3D Knowledge-based Segmentation Using Sparse Hierarchical Models : contribution and Applications in Medical Imaging

Salma Essafi

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

centre de recherche SACLAY - ÎLE-DE-FRANCE

PHD THESIS

to obtain the title of

Doctor of Ecole Centrale Paris

Specialty : Applied Mathematics

Presented by

Salma ESSAFI

3D Knowledge-based Segmentation Using Sparse & Hierarchical Models:

Contributions and Applications in Medical Imaging

prepared at Ecole Centrale de Paris, MAS Laboratory

Defended on May 12, 2010

Committee :

Chairman :	Alain RAHMOUNI	-	Henri Mondor Hospital
Reviewers :	Allen TANNENBAUM	-	Georgia Institute of Technology
	Herve DELINGETTE	-	INRIA - Sophia-Antipolis
Advisors :	Nikos PARAGIOS	-	Ecole Centrale de Paris
	Georg LANGS	-	Massachusetts Institute of Technology
Examiners :	Daniel RUECKERT	-	Imperial College London
	Gabor SZEKELY	-	ETH Zurich

2010ECAP009

Acknowledgements

It is a true pleasure for me to thank the people who have made this thesis possible.

First of all, I would like to thank my director, Nikos Paragios, for giving me the chance to accomplish my PhD thesis at Ecole Centrale Paris and for guiding me in this long journey. His extraordinary leadership and research expertise constantly pushed me to go beyond my limits.

I owe my deepest gratitude and respect to my supervisor Georg Langs who has shown his support in many different ways. Thanks for all our passionate discussions and brainstorming sessions about my thesis, for all the time dedicated to explain me the tiniest detail, for all the meticulous effort to improve every single paper and presentation. Thank you forever for giving me your generous support and invaluable guidance and inspiration all along the time it took to accomplish this thesis. Your apprenticeship and friendship are a privilege and a pleasure.

This thesis would not have been possible without the financial support from the AFM "Association Française contre les Myopathies“ under the DTI-MUSCLE project.

It has been an honor for me to have as committee members Prof. Alain Rahmouni, Prof. Daniel Rueckert, and Prof. Gábor Székely. Their presence has been gratefully acknowledged. A special thank to my thesis reviewers Prof. Allen Tannenbaum and Dr. Hervé Delingette for their feedback and their precious time dedicated to reviewing the manuscript despite their numerous constraints.

I am highly grateful for ground truth segmentation, and for all the priceless time Dr. Jean François Deux from Henry Mondor Hospital provided to us in order to be able to accomplish the MRI acquisitions, without which our methods and codes would have been meaningless.

I am indebted to my many unofficial thesis readers: Lilla Zollei, Mickael Bronstein as well as my dear office-mates.

In the spring of 2009, I had a very inspiring and fruitful stay at the CIR lab in the

Medical University of Vienna, mainly thanks to all the crew team and to the studious and friendly atmosphere there, a warm thank to Janina Patcsh, Eva Dittrich, Rene Donner, Alexander Valentinitsh, Ernst Schwartz and all the others! It has been an honor, too, to be able to spend two months at Yale University in the summer 2007, in the Biomedical lab of Professor James Duncan.

Many thanks to Sylvie Dervin, the shining star of our lab, who has always been present when I needed help, thanks Sissi. I am also grateful for Annie Glomeron's and Christine Biard's understanding and help concerning all the administrative issues.

I was very fortunate to have Aris Sotiras, Loic Simon, Olivier Teboul, Mickael Savineaud, Pagnaiotis Koutsourakis, Ahmed Besbes and Radhouène Neji as my office-mates. They have made my PhD journey at Centrale incredibly interesting and fun, we shared a lot of good memories starting with our running sessions to catch the last RER B at midnight, our nights spent in the office before critical deadlines, with their unfailing help during my (numerous) moving episodes in Paris, our different trips around the world (I particularly recall Easter in Greece...). I have learned a lot from them, and I am forever grateful for every moment that felt so hard to live, they were trying even harder to make me smile and make me see the bright side!

A warm thank you to my lab-mates at MAS, especially Katerina Gkirtzou, Chahoui Wang, Fabrice Michel, Regis Behmo, Noura Azzabou, Siddharth Shankar... the list is too long.

All along those 4 years it has been a real pleasure to know Charlotte & Olivier Juan. Their valuable friendship was a real gift to me, even after they have left the lab their advices and constant help were really precious to me.

To my wonderful companions during my internship at MAS, Ben Glocker, Kush Varshney, Christian Wachinger, and Lilla Zollei. I was so lucky to find them, even though the lab was quite intimidating in the beginning. With the 2006 World Championship tournament everything went well in the end.

I am also truly thankful for Professors Faouzi Ghorbal and Slim Mhiri, my undergraduate and master advisors, as well as for the members of my former lab GRIFT in Tunisia, who have introduced me to the exciting world of computer vision and kept encouraging me all the way long.

Living in Paris is one of the best experience ever. This has been possible thanks to my friends, Abderrahmen Mtibaa, Insaf Meliane, Dali Hassine, Hend Ennouri, Achraf

Madani, Danielle et Pierre Morisseau, Naouel Zitouni, Victoria Arnold and Taha Krichene. Moreover I would like to extend my thanks to my dearest friends back home in Tunisia, Ons Sakji, Omar Smaoui, Aicha Hbacha, Dali Maktouf, Wajdi Selmi, Dali Chermi, Amina Radhouane, Chiraz Jemour and Hela Fourati, who have always been a great source of moral support and with whom I shared endless moment of happiness.

I owe a tremendous amount of gratitude to my big and generous family, who never really understood what I am doing, but who was always proud of me! I would like to express my gratitude to my godfather Abdel, he has always been, and still, an invaluable source of support and guidance. For my sister Sonia and my brother Skander who are always making fun of me and how it is possible that I am still studying at this age! To my grand parents who are no longer here to witness this achievement, but whose love and trust is always and forever present to keep me going.

This thesis is dedicated to my beloved mother for all the sacrifices she has made for us, for her endless love and unconditional support with respect to all my aspirations.

To the love of my live, my husband Adel, without him I would not be what I am, and the achievement of this thesis would have never been possible... I am looking forward to all the coming adventures!

A la mémoire de mon père...

"La vie est une comédie, laquelle il n'importe combien elle soit longue, mais qu'elle soit bien jouée.“ François des Rues (1595)

Abstract

THE thesis is dedicated to three dimensional shape analysis and the segmentation of human skeletal muscles in the context of myopathies and their treatment. In particular, we study the local and global structural characteristics of muscles. The methodological focus of the thesis is to devise methods for the segmentation of muscles, the consistent localization of positions in the anatomy and the navigation within the muscle data across patients. Currently diagnosis and follow-up examinations during therapy of myopathies are typically performed by means of biopsy. This has several disadvantages: it is an invasive method, covers only a small muscle region, is mainly restricted to diagnostic purpose and is not suitable for follow-up evaluation. We develop the following methods to make the use of non-invasive imaging modalities such as MRI for a virtual biopsy possible: first, a novel approach to model shape variations that encodes sparsity, exploits geometric redundancy, and accounts for the different degrees of local variation and image support in data. It makes the modeling and localization of muscles possible, that exhibit sparsely distributed salient imaging features, and heterogeneous shape variability. Second, we extend the shape representation of 3D structures using diffusion wavelets. The proposed method can represent shape variation and exploits continuous inter-dependencies of arbitrary topology in the shape data. We then explore several approaches for the shape model search, and appearance representation based on boosting techniques and canonical correlation analysis. Last we present a robust diffusion wavelet technique that covers the integration of our two shape models approaches to finally get an enhanced sparse wavelet based method. We validate the approaches on two medical imaging data sets that represent the properties tackled by the approaches: T1 weighted MRI data of full calf muscles and computed tomography data of the left heart ventricle.

Key words: Segmentation, Shape Analysis, Sparsity, Diffusion Wavelet, MRI, Skeletal muscle, Myopathy.

Résumé

CETTE thèse est consacrée à la conception d'un système d'aide au diagnostic dédié au muscle squelletique humain. Au cours du premier volet de ce manuscrit nous proposons une nouvelle représentation basée sur les modèles parcimonieux dans le cadre de la segmentation d'Images de Résonances Magnétiques (IRM) T1 du muscle squelettique du mollet. Notre méthode Sparse Shape Model/ Modèle de Formes Parcimonieux (MFP), apprend un modèle statistique de formes et de textures locales annoté et réussit à en tirer une représentation réduite afin de reconstruire le mécanisme musculaire sur un exemple test. Dans la seconde partie du manuscrit, nous présentons une approche basée sur des ondelettes de diffusion pour la segmentation du muscle squelettique. Contrairement aux méthodes de l'état de l'art, notre approche au cours de la phase d'apprentissage permet à optimiser les coefficients des ondelettes, ainsi que leur nombres et leur positions. Le modèle prend en charge aussi bien les hiérarchies dans l'espace de recherche, que l'encodage des dépendances géométriques complexes et photométriques de la structure d'intérêt. Notre modélisation offre ainsi l'avantage de traiter des topologies arbitraires. L'évaluation expèrimentale a été effectué sur un ensemble de mollets acquises par un scanner IRM, ainsi qu'un ensemble d'images tomodensitométriques du ventricule gauche.

Mots Clés: Segmentation, Analyse de Formes, Modèles Parcimonieux, Ondelettes de diffusion, Imagerie par Résonnance Magnètique, Muscle squelettique, Myopathie.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

"A goal without a plan is just a wish."

Antoine de Saint-Exupery

1.1 Motivation

The human body has no longer any secrets for the medical experts, thanks to the growing variety of image acquisition techniques. Physicians have access to an incredible number of image acquisitions technologies that allow for the quantitative observation of anatomy, physiology and pathology. Such modalities call for novel methods to make optimal use of the information. The role of these techniques is not to replace the doctors but to provide them with specific and rich information for their decisions.

Data acquired by medical imaging modalities have indeed achieved a level of richness that needs computer based methods to extract relevant information in a consistent and efficient manner. Imaging techniques hence play a key role in diagnosis and follow-up of muscle disorders in conjunction with clinical examination, biological analysis and muscle biopsy.

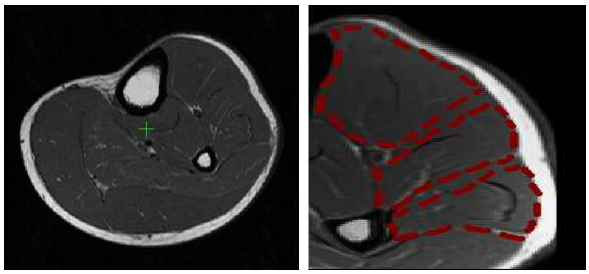
Myopathies are diseases that affect muscles connected to bones, called skeletal muscles, and produce weakening and atrophy of skeletal muscles, especially those closest to the center of the body resulting in reduced mobility. Such disorders affect a large percentage of the population with worldwide incidence of all inheritable myopathies being at about 14%.

Clinical follow-up as well as therapeutic trial evaluation are mainly based on functional tests and physiological measurement of muscle strength that are limited by the lack of sensitivity or poor reliability. Muscle tissue biopsy allows a precise microscopic myofiber count but this invasive method, covering a small muscle area, is mainly restricted to diagnostic purpose and is not suitable for follow-up evaluation.

Generally, diagnosis involves several outpatient tests to determine the type of myopathy like blood tests and more frequently electromyogram or still muscle tissue biopsy. Magnetic Resonance Imaging as well as Diffusion Tensor Magnetic Resonance Imaging are , on the other hand, techniques that allow to gather in-vivo measurements about the skeletal muscular diseases.

The scope of this thesis is to develop a framework for shape modelling and segmentation of the human skeletal muscles, and more particularly those affected by myopathy. The medical imaging community has focused on organs other than the calf muscle, for instance heart or brain tissues during the last decade. The segmentation of individual muscles within a muscle compound depicted with *Magnetic Resonance Imaging (MRI)* (see Figure. 1.1, 1.2) poses new challenges to automatic segmentation systems. Although

dominated by the global anatomy, muscle deformation exhibits mostly locally consistent behavior, precluding the use of global models. Muscle surfaces are only partially visible, while parts exhibit structures, that can change dramatically between patients, or during the course of follow-up examinations.



(a) (b)

Figure 1.1 — Calf MRI: (a) healthy (b) zoom on AT, EDL and PL muscle groups.

One of the most prominent processing methods in the medical image analysis is segmentation. The segmentation of images are necessary in different medical applications ranging from computer assisted surgery, study of anatomical structures , computer aided diagnosis and the monitoring of disease progress. The difficulty of segmentation mainly remains in the tremendous variability of objects and the variation in image quality. Additionally, there are a lot of complex properties, such as artifact and noise, corrupting the medical images and leading to the failure of common segmentation techniques.

Model-based segmentation and localization approaches are of prime interest, due to their ability to repeatably identify positions in the anatomy, to learn and to apply properties of large and representative populations. Of particular importance is the accurate representation of subtle local shape variation and the correct parametrization of the associated interdependencies between parts of the anatomical structures. This is a very tedious task to be performed manually, and thus has to be learned from the data, in order to avoid a bias that would compromise the model efficiency and accuracy. A second aspect that becomes prevalent when exploring large anatomical regions, is the highly heterogeneous nature of their appearance, and the existing local consistency across patients.

This thesis is part of the “DTI-Muscle” project, a collaboration between Ecole Centrale Paris, the French Association against Myopathy and Henry Mondor Hospital. Among the objectives of this project one can cite the use of a relatively recent non-invasive image modality to understand the global effect of different myopathies to mus-

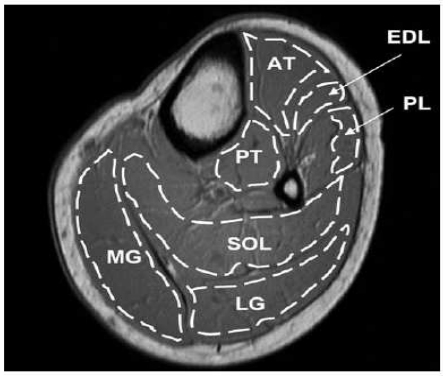


Figure 1.2 — Manual expert annotation: ROIs of the calf muscle

cles, the development of mathematical models capable of analyzing such data toward non-invasive diagnosis among the different myopathies as well as the exploitation of such modality to measure the reaction of muscles to treatment using different agents. Our contribution consists of developing the required techniques to use such an image modality towards modeling, understanding and diagnosing the muscular diseases.

The clinical goal is to devise a method to perform a virtual biopsy of muscles: the non-invasive analysis of the local and global structure of muscles and its changes during myopathies. The methodological prerequisites for such an automatic method are challenging: they involve the separation of different muscle types, the consistent localization of positions within low-contrast and only partially reliable data. The project thus not only contributes to the therapy, but also to the research in muscle diseases and their characteristics.

1.2 Contributions

The main contribution of this thesis are methods for the segmentation and localization of muscles in MRI data. On the one hand, the contributions explore modelling methodology that can cope with highly variable shape and sparsely distributed image information. On the other hand, we apply these methods to the specific question of muscle segmentation.

The contributions of this thesis include :

- A novel statistical shape model that aims to identify the challenging borders be-

tween muscle constituents, where the separation is not present along the entire boundary but rather on sparse locations only. Our approach introduces a robust version of the active shape models framework, with additional properties that adapt to sparse boundaries exploring how to construct sparse shape models that use only a subset of the available landmarks. The particular subset is optimally chosen based on, first, geometrical accuracy and influence on overall shape and second, the discriminative power of the local appearance model. We obtain thus a sparse model of objects or anatomical structures that takes the local statistical modeling and appearance behavior into account. We compare its reconstruction and search behavior with standard shape models, that neglect these properties, and represent objects regardless of the reliability and complexity of shape and texture behavior in the training population.

- A novel multi-scale shape representation and segmentation for anatomical structures using diffusion wavelet shape priors. The approach deals with complex and soft connectivity properties of objects by encoding their interdependencies with a diffusion kernel. The topology is learned from the training data instead of using a priori choices like e.g., a sphere, and represents the shape variation of structures with arbitrary topology by means of diffusion wavelets. The proposed model-based segmentation approach accounts for the systematic behavior of shape variation and image support in anatomical structures, with a parameterization that goes beyond pre-defined reference manifolds. For the parameterization of complex structures, it is worthwhile not to rely on a reference manifold with an a priori topology, but to learn the appropriate topology from the training data. For this, we have to determine the intrinsic topology a shape for which multiple examples are available, and have to encode this information in the shape model, and to use it in the representation and during the segmentation.

  Given a set of training shapes and their corresponding sets of landmarks, the goal is to construct a manifold on which the training shapes live and to model the shape variation on this manifold hierarchically using diffusion wavelets. Toward addressing segmentation, statistical learning at different levels of hierarchy (diffusion kernels) on the variations of the wavelet coefficients is performed. Due to the power of the basis function representation, conventional dimensionality reduction techniques using the orthomax criterion [Kaiser, 1958] lead to a very compact representation of the manifold. The approach has definitely the potential to be used in many other computer vision tasks where data live in a very high-dimension space.

The experiments curried out on both calf and Left ventricle muscle data indicate clear advantage of the proposed method over the competing ones.

- Different search strategies were introduced, for instance the use of our statistical wavelet-based model together with a GentleBoost classifier, or with Canonical Correlation analysis. The approach obtains an accurate delineation of partially visible surfaces and complex textures, that cannot be achieved with registration based methods. During the search, the hierarchical diffusion wavelet shape model is fitted to new data based on local appearance captured by the classifier. The method computes a local feature vector for every voxel and maps it via a GentleBoost classifier to a probability that the voxel belongs to a specific landmark in the object. The classifier is trained from the data set segmentations. The probabilistic output is constrained by the shape model. The mapping onto the diffusion wavelet coefficients space ensures valid results with regard to the training data. The result of this procedure is a probability for each voxel regarding its match to the structure to be segmented, conditioned on both local and global information. We report results on CT left heart ventricle data sets, that illustrate the impact of the soft parameterization, as well as the global classifier based search.

- A robust diffusion wavelet method that covers the integration of our two approaches to finally get an enhanced sparse wavelet based method. This association undoubtedly inherits the advantages of both relative techniques. The resulting technique is independent from the topology of the anatomical structure, and can represent complex geometric and photometric dependencies of the structure of interest. This leads to a fully automated segmentation system using 3D-shape models for anatomical data.

To summarize, our contribution in this work is three-fold:

- Exploit sparsity present in the data and get rid of the existing information redundancy.

- Learn the appropriate topology from the training data, and use a corresponding shape representation based on diffusion wavelets to model its variation.

- Combine different search strategies based on classifiers and local descriptors.

1.3 Thesis Roadmap

We introduce in Chapter. 2 the medical background of our studies that rely particularly on human skeletal anatomy. Anatomical diseases and characteristics of the calf muscle will be investigated all along this section to introduce the reader to the medical part of our work.

Chapter. 3 provides an overview of medical image segmentation as well as the gold standard techniques of statistical shape modeling that constitute the state of the art in the computer and medical vision community. We do also discuss various issues surrounding shape representation such as alignment and landmarks correspondence.

In Chapter. 4, we propose a model that uses statistical properties of shape and appearance in a manner different from standard shape models. A sparse model is obtained that takes the non-uniform distribution of image information and model complexity on the surface of anatomical structures into account. Our approach is not constrained to a single reference manifold, but can capture and model structures of arbitrary topology.

Chapter. 5 covers a 3D geometric shape model for anatomical structures using diffusion wavelets, and applies this to segmentation. The diffusion wavelet approach is first described, then its use to model shape is explained where shape variations are learned at multiple scales, hierarchically capturing the model parameters. Indeed the topology of the structure is represented by the diffusion kernel, learned from annotated training data. Finally, the application to segmentation based on a detection scheme is addressed.

In Chapter. 6 we discuss how the shape model obtained either through our sparse approach or the diffusion wavelet method could be exploited during different search strategies.

Chapter. 7 proposes a combination of the latter presented techniques, sparse model and diffusion wavelet, to generate an enhanced robust wavelet model.

At the end of each chapter, we present the experimental validation that we conducted to demonstrate the approach over medical data sets.

Finally, we draw the conclusion of all the theoretical and experimental work provided in this manuscript, while suggesting different opening research possible in the future.

CHAPTER

2 Medical Background

"The body is a big sagacity, a plurality with one sense, a war and a peace, a flock and a shepherd."

Friedrich Nietzsche

2.1 The Human Musculoskeletal System

2.1.1 The Skeletal Muscle Architecture

Starting from the first edition of Henry Gray's Anatomy of the Human Body in 1858, commonly known as *Gray's Anatomy*, which has still been considered as one of the landmark in the field up until the most recent edition published in 2008 [Gray, 2008], the human musculoskeletal system always found major interest among scientist and anatomist. Amongst the diverse clinical studies established about human skeletal architecture, we note significant consideration in muscle fiber [Narici and Maganaris, 2006], changes in muscle size and architecture following, modeling and analyzing the musculoskeletal system [Delp and Loan, 1995], skeletal muscle adaptations to disease states [Gea *et al.*, 2006]. The basics of skeletal muscle mechanics are described in [Herzog, 2000]. The human skeleton (see Fig. 2.1) of a healthy adult is made up of 206 bones. Surprisingly at birth babies have 300 bones, obviously we do not loose muscle but they simply fuse during growth.

The human skeletal system has four different functionalities serving our bodies; mainly providing movement, support, protection, and also blood cells. The movement of our body is made possible through a cooperation between bones, joints and ligaments on one hand with the muscles and tendons on the other hand: the support because it gives the shape, size and framework of the body; the protection thanks to what our internal vital organs can be kept safe from injuries; finally the bone cells of our body do rely on blood to keep them alive and to bring them necessary food and oxygen. This explains why the bones are able to mend in case they are broken, proves how the skeletal system is far from being just a group of bones, and how it is managing in many interesting ways to preserve our bodies strong and healthy.

Along with the recent considerable advancement in medical imaging technology, the wide range of muscle exploration is getting larger and larger. A part from imaging modalities development, diverse software and medical applications have been conceived in order to treat, exploit and of course display anatomical images, which could be CT, MRI, or even ultrasound. Certainly a joint effort and collaboration between mathematicians, engineers and of course doctors has to be established in order to make this progress possible.

During the muscle architectural analysis one can distinguish between three commonly

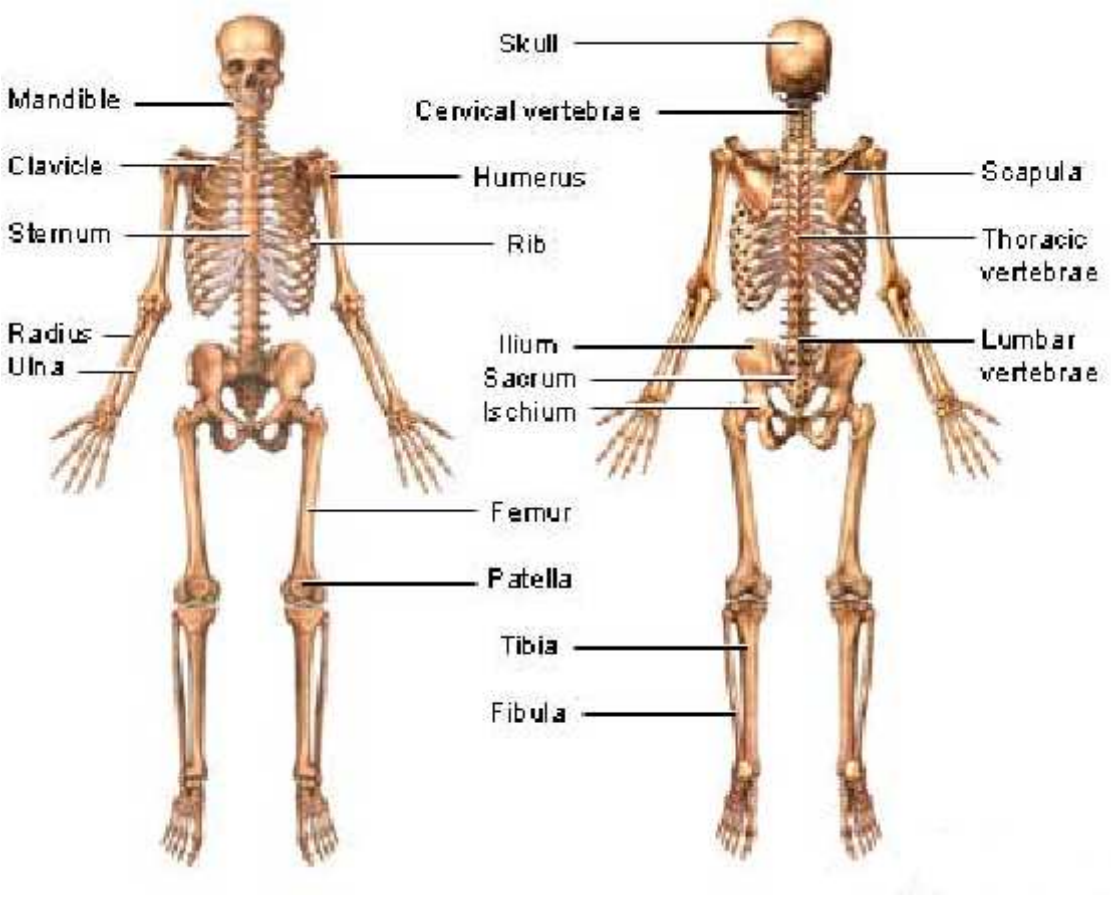


Figure 2.1 — The human musculoskeletal system (Copyright 2001 adam.com, Inc)

involved parameters which are ; Length (Lm), Fiber Length (Lf), Pennation Angle (e.g., the fiber angle relative to the force-generating axis, θ), and finally the Physiological Cross-Sectional Area (PCSA). Indeed quantitative studies in such fields get inspiration from micro-dissection of whole muscles.

We will focus our interest on the lower leg -Figure. 2.2- and more particularly on the calf muscle. In fact, the foot contains bones especially designed for weight-bearing. The latter represents a system of arches allowing for the foot to support much weight.

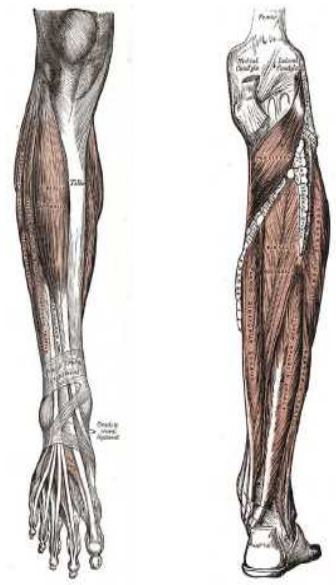


Figure 2.2 — The faciae of the lower leg

2.1.2 Calf Muscle Anatomy

From a physiological point of view, the calf belongs to the group of *"peripherical muscles"*, which are implicated in the limb movement. Indeed, the muscles of the calf act as the chief extensors of the foot at the ankle-joint. They exhibit significant power and are extremely vital in any every day movement.

In Fig. 2.3, the regions of interests (ROIs) are depicted in a calf slice [Basser and Pierpaoli, 1998], using an individual high contrast image and the corresponding labels for seven different muscles. The different muscles analyzed were, respectively, the soleus (SOL), lateral gastrocnemius (LG), medial gastrocnemius (MG), posterior tibialis (PT), anterior tibialis (AT), extensor digitorum longus (EDL), and the peroneus longus (PL).

When examining the calf anatomy, we find that the Gastrocnemius represents the most superficial muscle, composing the greater part of the calf, it also supplies to flex the femur upon the tibia, supported by the Popliteus. Located just right in front of the Gastrocnemius, there is the Soleus that is a broad flat muscle . For example, in walking, our whole body is supported on the raised foot, while during standing the Soleus insures the body from falling forward and maintains its stability. One also has to keep in mind that skeletal muscles are very sensitive and even extremely plastic [Gea *et al.*, 2006].

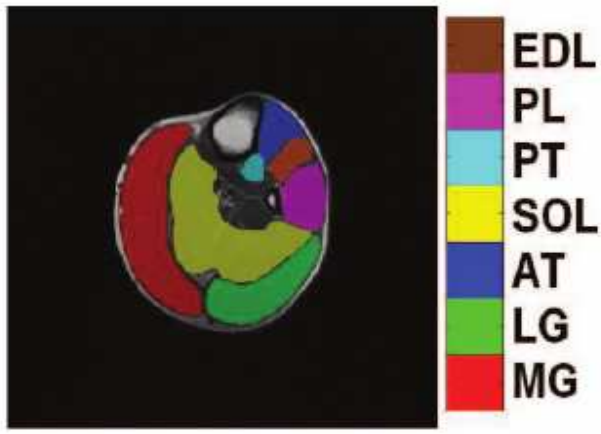


Figure 2.3 — T1-weighted axial slice depicting a supervised segmentation of the calf muscle. The seven different muscles are respectively; the soleus (SOL), lateral gastrocnemius (LG), medial gastrocnemius (MG), posterior tibialis (PT), anterior tibialis (AT), extensor digitorum longus (EDL), and the peroneus longus (PL).

However, muscle dysfunction could occur and trouble the human skeletal strength and endurance, resulting in the inability to maintain any kind of a task.

The boundaries between skin and muscle seem to vanish, which makes the segmentation procedure more and more complicated for an expert as well as for an automatic algorithm. Additionally the calf poses several challenges to the state of the art methodologies, e.g. the very subtle differentiation between muscles, and the un-even distribution of reliable image information.

The information provided in this section is deducted to a large extent from the most recent Gray's Anatomy edition [Gray, 2008]. For more extended details, we refer the reader to this estimable source.

2.2 Myopathy

Myopathies are diseases that affect the muscle system, and lead to a severe deterioration of the motoric abilities. These pathologies affect 4 to 6% of the population, i.e 25 to 30 million Europeans. Diagnosis as well as follow up for a given therapeutic strategy are often performed through biopsy. Magnetic resonance imaging (MRI) allows a non-invasive observation of the muscle fibers, their texture, and their global structure. It

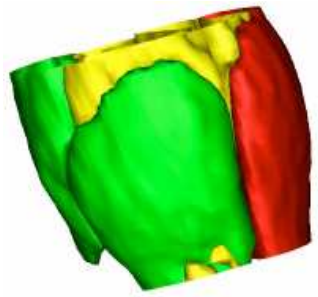
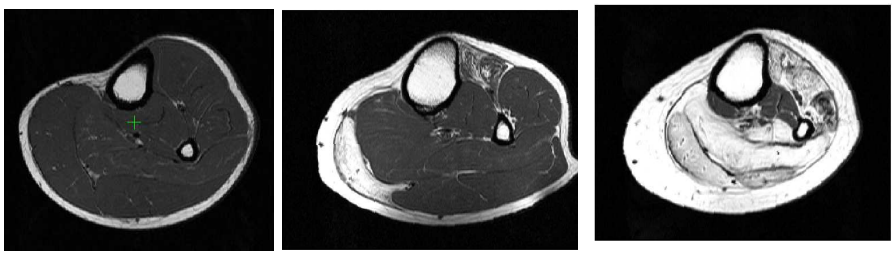


Figure 2.4 — Ground truth of the calf muscle.

has the potential to replace biopsies (by e.g., diffusion tensor imaging (DTI)) and has the advantage of encompassing a section of the entire leg as opposed to small local samples. This enables the analysis of local properties, as well as understanding the global structural change of muscles affected by a disease. A crucial first step in this analysis is the accurate segmentation of individual muscles.

First, myopathy symptoms came along with childhood and adolescence period. They had to be treated early and fast to prevent the disease progression and harm to the body. Indeed, in several dystrophies the heart can be seriously affected and this may even lead to death in some cases. One can characterize the muscular dystrophies, which are inherited myogenic disorders, by progressive muscle wasting and weakness of variable distribution and severity.



(a) (b) (c)

Figure 2.5 — Calf MRI: (a) healthy case (b) and (c) unhealthy case where the fat in white is spreading all over the muscle.

Acquired Myopathies	Polymyositis and Dermatomyositis Systemic Lupus Erythematosus Rheumatoid Arthritis Inflammatoy bowel disease myopathies Endocrine myopathies Rhabdomyolysis Steroid and other drug myopathies Electrolite disorders
Inherited Myopathies	Muscle Dystrophies Duchenne muscular dystrophy Steinert myotonic dystrophy Facio-scapulo-humeral muscular dystrophy Limb-girdle muscular dystrophy Congenital myopathies Metabolic myopathies Mitochondrial myopathies Acid maltase defficiency
Neurological-neuromuscular junction disorders	Amyotrophic lateral sclerosis Multiple sclerosis Guillain-Barre syndrome Parkinson disease Myasthenia gravis

Table 2.1 — Review of most predominant myopathic diseases and neuromuscular junction disorders [Gea *et al.*, 2006].

2.2.1 Neuromuscular Diseases Classification

These disorders can be classified in several groups, including congenital forms, in accordance with the more distinct distribution of muscle weakness: Duchenne and Becker; Emery-Dreifuss; distal; facioscapulo-humeral; oculopharyngeal; and limb-girdle which constitutes the most heterogeneous class. Among the disorders listed in Tab. 2.1, we will mainly detail the two most common ones, Duchenne Muscular Dystrophy and Steinert Myotonic Dystrophy. More exhaustive description of the rest of the list can be found in

[Gea *et al.*, 2006] and references therein.

The Duchenne Muscular Dystrophy (DMD) also known as *Pseudohypertrophic Muscular Dystrophy* represents the most prevailing and also the most severe dystrophic myopathy. For the historical part, it is named after Dr. Duchenne de Boulogne who worked in Paris in the mid-19th century, and who was one of the first doctors to show interest in muscular dystrophies.

Actually, this recessively inherited disorder is at the origin of skeletal muscle progressive degeneration. The genetic abnormality is caused by the defective gene encoded for the Dystrophin, the latter represents actually a protein for the muscle. Biologically speaking, Dystrophin is a protein located close to plasmolemma and linked to it by glycoproteins, that are respectively connected to laminin, as laminin is to the basal membrane [Emery, 1993]. Dystrophin fortifies the plasmolemma in such a way that it can be more resistant facing the mechanical forces coming from contraction-relaxation cycles. Once dystrophin is absent, injury is more likely to happen.

Male are more likely to get this specific kind of distrophy, whereas women act just like carriers as explained in Fig. 2.6. This is due to the severe recessive X-linked characteristic of the disease and also to the absence of descendants from affected men, analogously women always receive at least one disease-free X chromosome.

As for the symptoms, they come along first at around the age of five years. One of the first trouble facing the patients is walking difficulties as well as frequent falls followed by some change in the appearance. Those problems are due to the early hip and lower limb muscle implication. Muscle weakness first affects feet, fronts of thighs, hips, afterward more muscles get affected such as abdomen, shoulders and also vertebral column. Furthermore the fat infiltration of muscles generates a deformation in the limbs Starting from the age of 12 years old, patients are not able to walk and thus need a wheelchair. Along with respiratory even cardiac muscles start to represent a handicap. Without adequate support treatments, patients die on average between 20 and 25 years old either from heart failure or pneumonia.

Myotonic Muscular Dystrophy (MMD) also known as the Steinert disease represents the second most prevealant dystrophy after the DMD -Duchenne Muscular Dystrophy. Contrary to the DMD, the Steinert Myotonic Distrophy can happen at any age, where the patient starts wasting his muscles, and ends up with heart defection issues. It is a

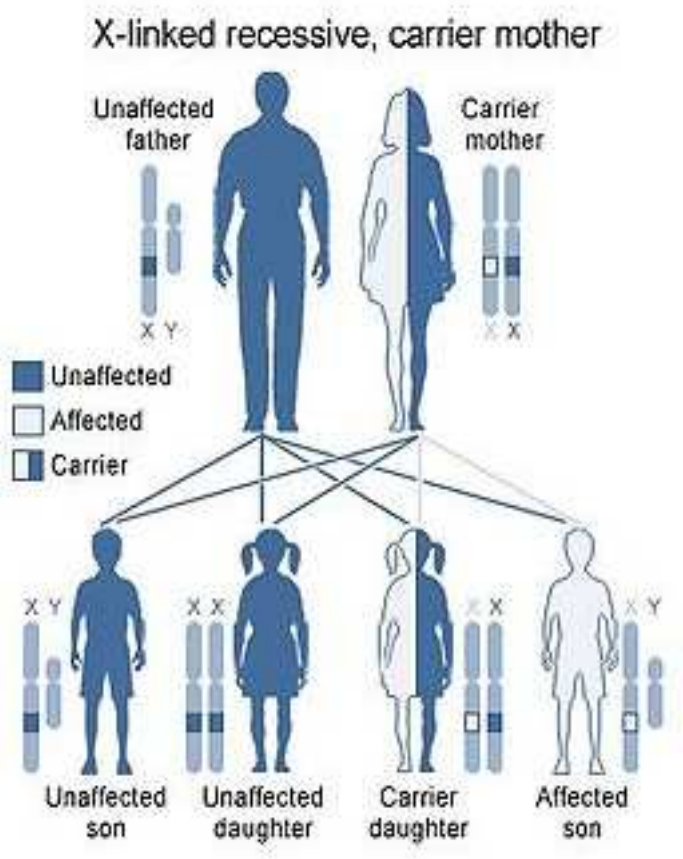


Figure 2.6 — Duchenne Muscular Distrophy: how the inheritence works (@Copyright. US National Library of Medicine)

slowly progressing and chronic disease, where the term “Myotonic” is the adjective for ” Myotonia", the inability to relax muscles at will.

As for symptoms severity, they can vary considerably, even within the same family. Nevertheless the substantial difference resides between the congenital cases and the cases beginning during adolescence/childhood. In fact, newborn infants are affected by severe muscle weakness, that could even lead to life-threatening problems requiring intensive care.

There are two types of Myotonic Muscular Dystrophy which are designated type 1 and type 2. The clinical features of type 1 tend to be more severe resembling to those of type 1, the two types are caused by mutations in different genes.

In terms of genetics, the gene which causes the dystrophy of type 1 is known as DMPK and is situated in the long arm of chromosome 19. The gene is coded for myotonic dystrophy protein kinase, which is a protein expressed primarily in skeletal muscle.

2.2.2 Possible Treatments

Due to the notable advances in gene manipulation, the genes and their corresponding protein products responsible for these disorders can now be detected. This identification represents a crucial step to corroborate an accurate diagnosis for not only the patients but also for prenatal diagnosis.

Unfortunately, there is, as yet, no way of greatly affecting the long-term course of any of these disorders. Nevertheless, significant progress in several disciplines (e.g., gene manipulation, stem-cell therapy) offer cautious optimism for finding an effective treatment in the not-too-distant future.

In the particular case of muscular dystrophy, we distinguish novel experimental treatments and on-going research [Cossu and Clemens, 2001] that can be subdivided in three major classes:

- Gene therapy: generation of new viral vectors more effective in transferring genetic material from one bacterial cell to another within adult muscle fibers.

- Novel pharmacological approaches: how to recognize the molecules responsible for the skipping of the mutated axon, i.e., the sequence in the DNA or its RNA transcript.

- Cell therapy: a recent technique based on stem cell biology and more particularly on stem (progenitor) cells transplantation. Yet the medical advancement and knowledge remains minor, and has to be deepened in order to establish a more significant clinical protocol for such neuromuscular disorders.

2.3 Medical Image Modalities

The last few decades show a rising attention to modern non-invasive techniques. Researchers and physiologists are able to investigate the human body, so the latter has no longer any secret to keep.

Starting from the work of the Nobel-prize winner A.V. Hill [Bassett, 2002], who contributed a pioneering insight to the field of skeletal muscle imaging. The in vivo exploration of anatomical structures has finally become possible.

In this section, we provide an overview of the imaging modalities while highlighting the gold standard technique, e.g. Magnetic Resonance Imaging, as far as the muscle state diagnosis is concerned.

2.3.1 Acquisitions Modalities Overview

2.3.1.1 Diffusion Tensor Imaging (DTI)

Diffusion Tensor Magnetic Resonance Imaging (DTI) is a technique that allows to measure the random motion of water molecules in biological tissues in vivo such as the white mater of the brain where it has been shown to allow non-invasive mapping of connectivity. Myofibers refer to anatomical structures where the propagation/diffusion of water could lead to a complete understanding of the muscle structure. Such local and global structure is altered when muscular diseases are present. One has to account for the highly sparse data of such a modality (capturing diffusion in a limited number of directions), the presence of strong noise on the acquisition model, the extraction of the muscle fibers from isolated measures, and the understanding of the global muscle structure through the statistical characterization of these fibers. Furthermore, we would like to correlate DTI results with morphometric data resulting from myofiber examination by microscopic histological study of the same muscle. The objective of this project is the development of a novel quantitative method for in vivo muscle imaging (DTI-muscle) leading to "near virtual muscle histology". DTI-muscle may offer a new reliable non-invasive approach allowing quantification of myofibers in the setting of pharmaceutical drug evaluation as well as for gene and cell therapy clinical trials.

As demonstrated in [Bihan *et al.*, 2001], DTI inherits its power from its own concept, as during their random, diffusion-driven displacements molecules probe tissue structure

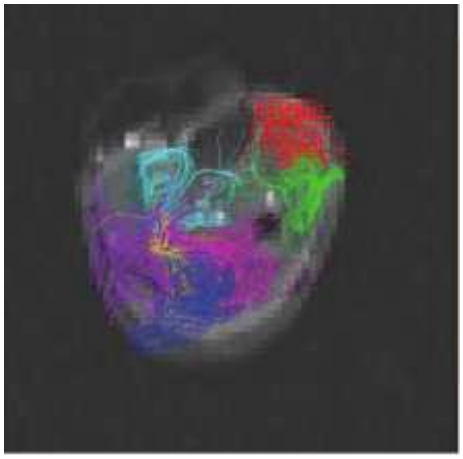


Figure 2.7 — Axial view of tractography classification for a healthy case [Neji *et al.*, 2009]

at a microscopic level far beyond the usual image resolution. Thanks to this acquisition technique, researchers are well equipped to specify and work on the diffusion anisotropy effects as well as tissue micro-structure. Another statistical characteristic one can observe in the diffusion MRI image voxel, is the movement of water molecules distribution inside the voxel. Scientific interest and understanding of DTI have steadily mounted, to gain access to diffusion tensor and especially to exploit fiber tracking in brain or more recently in the skeletal muscle.

Previous DTI studies in medical imaging [Bihan *et al.*, 2001] addressed more anatomical structures like brain white matter, tongue or still the cardiac muscle. As for the skeletal muscle, in [Galban *et al.*, 2005] the authors established how DTI is capable of separating the different variations in diffusive properties between different muscles. A more recent work [Neji *et al.*, 2009] handles the problem of calf muscle segmentation with respect to DTI modality where a novel manifold-based clustering approach for the classification of fibers is developed (see Fig. 2.7).

Computed Tomography (CT) The idea behind CT acquisition is to combine the digital computer together with a rotating x-ray device in order to generate detailed cross sectional images of the muscle slices, more especially bony structures of the foot or the ankle (see Fig. 2.8). CT helps to delineate the structures corresponding to these anatomical structures for 3D visualization, and it provides an accurate image of bones, soft tissue and blood vessels all at the same time. This in-vivo medical imaging technique

is commonly required in the case of fractures, infections, degenerative and rheumatoid arthritis, and more serious issues such as bone tumors, etc..

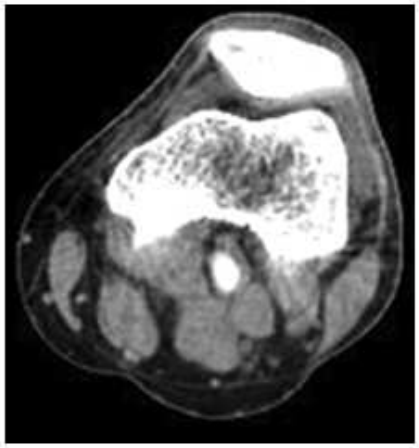


Figure 2.8 — Axial view of CT angiography depicting aberrant accessory muscle slip of medial head of right *Gastrocnemius* muscle [`American Journal of Roentgenology`]

Among the advantages of Computed tomography over other x-ray modalities one can cite how it is able to clearly expose the shape and precise location of soft tissues and bones in any slice of the foot/calf muscle. CT scans help doctors distinguish between a simple cyst and a solid tumor and any involvement of the bone. Roughly speaking, CT scanning can be considered as more rigorous than conventional x-ray in determining the stage (extent) of some bone tumors, and then it can guide the doctors in their decisions concerning the treatment. However, there is a remaining risk that it can lead to cancer due to the absorption of the radiation by the body in the case of excessive exposure to radiation, and it is also forbidden to some patients, like pregnant women. Indeed there is a high supposition of a linear relationship between radiation dose and cancer risk.

Ultrasound (US) The non-invasive and real-time method of ultrasonic ecography is devoted to measure muscle architecture, such as fiber bundle length, muscle thickness, and angles of pennation [Koryak, 2008]. It is considered to be a highly informative and available technique for assessing the architecture of human skeletal muscle. The fields of use of ultrasound scanning are varied, that goes from medicine, sports as well as physiology of aging. The relevant use of this technique relies on its accurate ability to assess muscle relationships with their functions, without assigning the patient to undergo painful procedures as muscle biopsy. More generally, this method is dedicated for the

heart and fetus imaging. Ultrasound uses high frequency sound waves and their echoes produce bright spots in the image between two different types of tissues.

Originally, humanskeletal architecture was studied and investigated starting from cadaver specimens, however, thanks to recent ultrasound studies [Fukunaga *et al.*, 1997] [Koryak, 2008], physicians are able to carry those study on human body, which leads to more rigorous and correct conclusions.

Comparing to other technique, like MRI or still tomography, ultrasound presents two advantages: being less expensive from the clinical aspect on one hand, and allowing physiologists to estimate muscle structural changes under functional stress on the other hand. Nevertheless, ultrasound does also suffer from some inconveniences, mainly low image quality, presence of geometric distortions caused by wave velocity, and even presence of speckle due to wave inference.

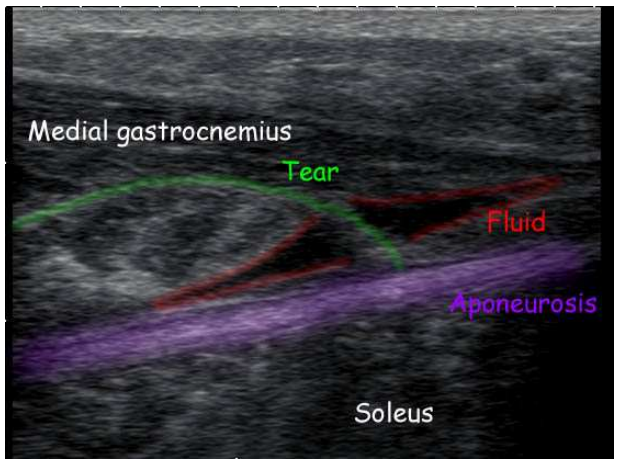


Figure 2.9 — Ultrasound estimation of a calf muscle tear.

Regarding muscular diseases, ultrasound (US) images have gradually lost their significance and they have been progressively replaced by Comuted Tomography (CT) or/and Magnetic Resonance Imaging. Nevertheless, due to its portability, low cost and lack of ionizing radiation, the ultrasound technique may be useful in some experiments. CT on the other hand is able to provide a fast and extensive evaluation of muscle anatomy and can detect fatty infiltration. Nevertheless, exposition to ionizing radiation limits the use of CT for follow-up of patients.

MRI has demonstrated significant advantages compared to the two former modalities. MRI imaging has so far proved to be a more accurate technique to evaluate and

follow-up muscle disorders [Filler *et al.*, 2004]. The major attraction concerning MRI is most probably the absence of ionizing radiation, which allows multiple examinations and evaluation of muscle disease progression.

2.3.2 Magnetic Resonance Imaging (*MRI*)

Since the introduction of Medical Resonance Imaging, i.e., approximately 30 years ago, the role of this volume imaging technique both in research and clinical practice is getting more and more considerable. Contrary to CT, MRI does not utilize ionizing radiation, instead, it employs a powerful magnetic field to align the nuclear magnetization of hydrogen atoms in water in the body. Indeed, the principles of nuclear magnetic resonance (NMR) represents the basis of Magnetic resonance imaging (MRI). However, to avoid any kind of negative connotations relative to the nuclear field, the term nuclear was finally abandoned in late 1970'.

As the human body mainly constits of water, which means two hydrogen nuclei or protons, when it undergoes an MRI acquisition, it is placed in a powerful and uniform magnetic field. Consequently the nuclei magnetic moments are aligned in parallel or anti-parallel way to the magnetic field. The moment rotates then around the axis of the field, with a certain frequency, known as the Larmor frequency, that relies both on the type of nucleus and on the strength of the magnetic field. The reaction of the tissues relies not only on the proton density, but also on the way that protons recover their initial/resting state position after the initial RF pulse (Radio Frequency).

Two parameters relative to the configuration of an MRI acquisition are repetition time (TR) and echo time (TE). TR is relevant to the period between two consecutive RF, while TE is the time separating the first RF and the echo.

Interest about MRI relies on how we could learn about their environment thanks to the energy emitted, when the excited nuclei relax and realign. We have two ways of measuring the relaxation time: T1-longitudinal relaxation- and T2 -transversal relaxation-, explained as following;

T1 represents the necessary time for the nuclei realignment, using a gradient echo (GRE) with long TE and long TR.

T2 assigned as the loss of phase coherence in the transverse plane, while utilizing spin echo (SE).

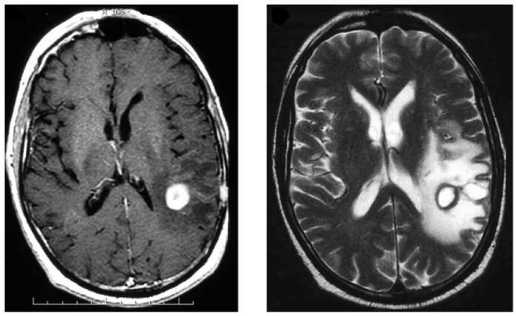


Figure 2.10 — T1 and T2 weighted MRI acquisitions of the brain.

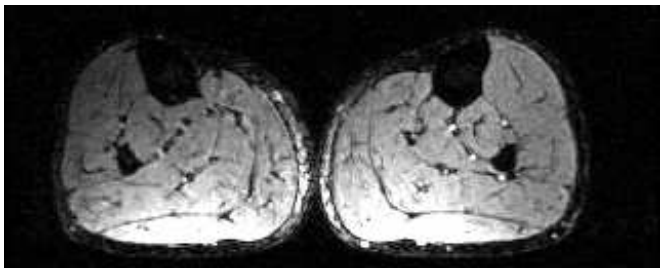


Figure 2.11 — Axial view of T2*-weighted image of the calf

T2* is an extension of the T2-weighted, where local magnetic field inhomogeneities are not equilibrated (compensated) as depicted in Fig. 2.11.

More generally, T1 and T2 -see Figure. 2.10- modalities are often used in clinical practice. The variation over the scanner parameters produces a contrast between different types of body tissue, so that researchers can detect diseased tissue, including tumors. Still this kind of exam is considered as expensive compared to other exams, but benefits from the fact that it is commonly known and admitted as being non-harmful to the body.

The excellent soft tissue resolution of MRI is superior to that of both CT and US. Thus, MRI is the most pertinent technique to detect fatty infiltration, edema -abnormal accumulation of fluid beneath the skin or in one or more cavities of the body-, and fibrosis -the formation or development of excess fibrous connective tissue in an organ. In the case of fatty infiltration, muscle signal is increased on T1-weighted spin echo sequences. Edema is a predominant finding in inflammatory disease and appears as an area of high signal intensity on T2-weighted spin echo sequences. Fibrosis of muscle tissue is recognizable only on MRI where it appears as a high signal area on T1-weighted spin

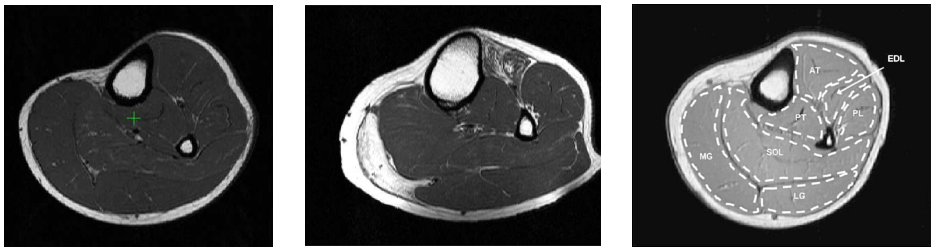
echo sequences after gadolinium administration. MRI provides key qualitative informations about muscle disorders. However, a reliable quantification method of the muscle structure remains to be obtained for precise follow up of patients and in the setting of evaluation of clinical trials aiming at increasing the muscle mass by pharmacologic approach or stem cells engraftment.

With regard to the the calf muscle issue, the Magnetic Resonance Imaging modality offers us the opportunity to investigate the muscle in a non-invasive way, and to discriminate tissues due to its high contrast (See Fig. 2.5). Besides high resolution image data sets are required to process to improve volume measurements. Hence, this goal is reached through 3D imaging procedures instead of 2D ones. Likely, recently the understanding and development of fast imaging sequences and muscle disorders have steadily mounted which grant large data acquisition in reasonable amounts of time as well as concluding investigations.

A more detailed description of the physical aspect could be found in [Macovski, 1983]. There is still a wide variety of other specialized MRI scans, such as Magnetic resonance angiography, Susceptibility weighted imaging (SWI), or Functional MRI.

2.4 Medical Problem Statement and Goals

In medical image analysis, often the assumption of consistent global regions statistics is violated within organs. The work was motivated by the highly heterogeneous distribution of visual information in muscle MRI data (Fig. 2.12). Standard segmentation methods fail, due to the ambiguous texture and the sparse distribution of salient image information within the leg. The majority of work on anatomy segmentation is focused on brain, liver or heart data [McInerney and Terzopoulos, 1995; Bizais *et al.*, 1995; Butz *et al.*, 2003]. They are often based either on the detection of prominent edges between organs or on the separation of intensities due to the reflectance properties of the different tissues in the human body. The segmentation of individual muscles within a muscle compound poses new challenges to automatic segmentation systems. The sparse distribution of regions where image information allows for a reliable separation of neighboring substructures, makes the use of prior shape knowledge mandatory, and motivates the development of algorithms, that make optimal use of statistical models of shape and appearance acquired during a training phase.



(a) (b) (c)

Figure 2.12 — MRI data of calf muscles: (a) healthy (b) and unhealthy case, (c) manual expert annotation of individual muscles.

In Fig 2.12, magnetic resonance imaging (MRI) slices of the human calf are depicted [Galban *et al.*, 2005]. The distribution of reliable image information at the boundaries between individual muscles is un-even, and parts can only be estimated from prior information about shape. Since this distribution is fairly consistent over a population, it can be learned, and integrated in a sparse model that makes optimal use of both the shape prior and the image information.

This thesis is therefore motivated by our belief that segmenting human skeletal muscles is attainable by exploiting the redundancy and the local image support present within the image. By properly considering the challenging features of the image when performing model analysis, we aim at significantly improving the application performance and propose solutions to the ill-proposed problem of calf muscle segmentation. Indeed, not much work has been devoted to the skeletal issue, which brings another intriguing part to our mission.

We will start by focusing on the shape modelling part, and on the major question of landmark generation and correspondance. The main issue afterward is to study the sparsity of these landmarks. We propose a model for joint shape modeling and segmentation. By leveraging the tools related to differential geometry (diffusion maps, diffusion distance..) and diffusion wavelets, we are able to handle the challenging medical data set and retrieve accurate results.

The notion of multiclass inherited from diffusion wavelets will also be exploited to analyze and decompose our data set into various levels. Such a decomposition will help us to discover different information content of the shape depending on the scale, the latter information is not necessarily available from looking at the original image volume.

2.5 Summary

In this chapter, we presented the medical background of the thesis, so that the reader gets familiar with a useful set of vocabulary terms related to Myopathy. To recapitulate, the most significant challenges of our study area are the following; first, the complexity of the calf anatomic shape; second the partial visibility compounds of this muscle in T1 MRI slice ground truth; and last the high variability in the distribution of image information. The aim of this study is to develop new mathematical models from MRI data in order to improve our understanding of muscular diseases. Before introducing our shape represention technique, we give a review of the state of the art medical image segmentation approaches in the upcoming chapter.

CHAPTER

3 Background and State of the Art

"Science is facts; just as houses are made of stones, so is science made of facts; but a pile of stones is not a house and a collection of facts is not necessarily science."

Henri Poincare

3.1 Medical Image Segmentation

3.1.1 General Overview

Segmentation is a fundamental problem in image processing, medical image analysis and computer vision. Generally one would like to create a partition of the image in regions with similar features, and thus subdivise data into a number of disjoint regions according to their appearance properties. Such a problem is difficult, since the individual voxel values are not sufficient information for correct segmentation. Especially medical imaging data is ambiguous and we typically need to integrate a priori knowledge about shape and appearance of anatomical structures to successfully segment them.

The integration of a prior knowledge about the anatomical structures is the focus of a large body of current research. In the following we discuss several families of approaches that are related to our methodology. We will point out their properties and limitations in the context of our segmentation task.

The capturing of the a priori knowledge is a difficult task, and methods range from approaches that assume explicit properties (e.g. elasticity, smoothness), to algorithms that use information obtained during a typically supervised training period. Supervised segmentation and classification represent a hard and time consuming task, and with the huge amount of medical data available this task becomes even more laborious.

Throughout the literature we can distinguish model-free and model-based ones. Model-free methods make no assumption on the geometric properties of the region of interest. Model-based methods introduce certain assumptions on the space of allowable solutions -*priors*. These approaches are useful in the context of medical image analysis where variations of anatomical structures are constrained by the anatomy, while at the same time pose and view-point variation are taken into account. We can illustrate schematically the segmentation problem by the Figure .

State of the art segmentation methods mainly rely on a clearly defined topology, and an object boundary characterized by salient features (e.g, edges), where the majority of work on anatomy segmentation is focused on brain, liver or heart data [Duncan and Ayache, 2000]. In the medical and imaging community, MRI is regarded as the method of choice concerning in-vivo volume measurements thanks to its high sensitivity regarding soft tissue and its non-invasive nature.

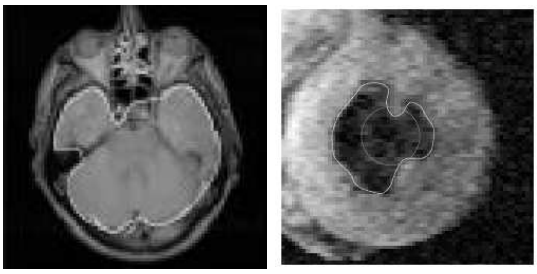
3.1.2 A Variety of Medical Segmentation Techniques

Boundary-based methods are seeking the optimal instance of the model guided by image-discontinuities, region-driven approaches aim to separate the global intensity characteristics of the object from the background, while feature-driven methods seek to learn patterns of support in the image. One can refer to a number of surveys published for snakes [Yezzi *et al.*, 1997], deformable models [McInerney and Terzopoulos, 1996], markov random field-driven [Tu and Zhu, 2002], active appearance models [Cootes *et al.*, 2001], minimal paths and fast marching techniques [Deschamps and Cohen, 2001], etc.

Deformable Contours and Surfaces Widely called upon in computer vision and pattern recognition, the deformable models are curves or surfaces defined within an image domain that can move under the influence of internal forces [Xu and Prince, 1998], [Staib and Duncan, 1996]. Thses models gained more attention since the publication of [Kass *et al.*, 1988]. A growing number of papers related to 3D deformable models [Terzopoulos *et al.*, 1988] [McInerney and Terzopoulos, 1996] [O'Donnell *et al.*, 1998] have been published in the literature. As a matter of fact the use of those models is becoming more and more common in medical imaging field.

Shape Models and Descriptors In [Staib and Duncan, 1992] the authors utilize elliptic Fourier descriptors in order to depict boundary templates. Free-form as well as parametric deformable models are among the key processing techniques in medical imaging segmentation. On one hand, one can find the Active Contours or Snakes as the most famous free-form deformable models and Active Shape and Statistical Shape modeling as the most well-known parametric deformable models. For a more exhaustive review of free-form deformable models in both 2D and 3D and their use in medical image analysis we refer the reader to [McInerney and Terzopoulos, 1996] and [Xu and Prince, 1998].

Nevertheless, elastic models suffer from some limitations in the way that they are often too flexible, and can be trapped by misleading edges or by edges adjacent to the structure of interest, converging to a suboptimal solution. Another drawback consists in the initialization requirement near the final solution, involving manual intervention of the user. Some papers, though, propose some solutions to tackle these problems like in [Kelemen *et al.*, 1999], [Székely *et al.*, 1996].



(a) (b)

Figure 3.1 — Examples of deformable segmentation images both for the brain case (a) and the left ventricle case (b)

Finite Element Finite element methods (FEM) are explored in [Gourret *et al.*, 1989], whereas [Teran *et al.*, 2003] exploit finite volume methods (FVM), which seems to be more intuitive than the finite element method (FEM), since they want to rely on a geometrical rather than a variational framework. Indeed B-spline solids are used in the latter paper in order to model fiber directions, and the muscle activation levels are derived from key frame animations. Another approach was presented by Cohen et al. in [Cohen and Cohen, 1993] where the authors generalize the ballon model -already introduced in [Cohen, 1991]- as a 3D deformable surface

Graph-based Technique Several graph-based approaches have been developed in the last decade, including the Normalized Cuts by [Shi and Malik, 2000], or the graph-cut segmentation proposed by [Boykov and Jolly, 2001]. A growing number of publications in vision use graph-based approaches for image segmentation, for instance one can cite [Veksler, 2000] and [Grady and Funka-Lea, 2004].

Indeed the tremendous development of efficient energy minimization algorithms has revolutionized the field of computer vision [Szeliski *et al.*, 2006], with methods based on Markov Random Fields (MRF) which have been used in medical imaging problem such as registration [Glocker *et al.*, 2008].

Summary Freely deformable models involve deformation algorithms that do not take into consideration the learning stage of shape variability constraints. Nontheless one has to keep in mind that freely-deformable models can be used to represent certain shapes, but the stabilizing energies are mainly based on general smoothness properties and not

on statistical consideration [Heimann and Meinzer, 2009]. More generally, conventional low-level techniques do not implicate inherent a-priori information, and hence suffer from lack of stability facing local image artifacts and perturbations present in the image. Consequently we will focus more on the model-based segmentation approaches, and ideally a shape representation that encodes variations at multiple scales, providing a large set of features for shape analysis.

3.2 Statistical Shape Models

Knowledge-based methods aim to satisfy such a demand where one first intends to determine some prior density on the space of solutions and then constrains the solution in a new image from this density. That way an accurate organ segmentation would offer precise measurements, simplify visualization and allow for a more accurate diagnosis.

Statistical models of shape and appearance variation form a basis for the work presented in this thesis. In the following we will outline the state of the art, and point out its capabilities, and limitations. Modeling shape variation is a well studied problem, where there are two critical components; the choice of shape representation and the construction of the prior manifold. For instance Point distribution or landmark-based models [Cootes *et al.*, 1995], implicit representations [Rousson and Cremers, 2005], triangulated surfaces [Welch and Witkin, 1994], spherical wavelet representations [Nain *et al.*, 2007] and statistical atlases [Rueckert *et al.*, 2003] are examples of shape/surface representations. Besides shape and appearance models have been used in different contexts, like face modeling [Edwards *et al.*, 1998], [Matthews and Baker, 2004], tracking [Dornaika and Ahlberg, 2004], studying human behavior [Johnson *et al.*, 1998], and medical imaging tasks.

Given the shape representation, the prior manifold can either be a subspace or a probability density function. In the first case, the space of solutions is often represented using a linear combination of a set of basis functions modeling the variations of the training examples. Linear sub-spaces, determined either through Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), or Non Negative Matrix Factorization (NNMF), are methods being used to determine these subspaces. Towards dealing with high amounts of training data, numerous provisions were considered such as Kernel-PCA prior [Cremers *et al.*, 2002], or sparse models, either through implicit [Florin *et al.*, 2007] or through explicit surface representations [Essafi *et al.*, 2008]. In the second

case, simple Gaussian densities, mixture of Gaussians, non parametric densities [Taron *et al.*, 2009] as well as manifold learning and embedding [Charpiat *et al.*, 2005] were considered.

In [Rogers and Graham, 2001] the authors describe how to model present features, which deals precisely with objects with missing data. Still in the same context, the authors in [Rogers and Graham, 2002] use a weighted least squares to update the shape parameters. In this case, landmarks with distinctive muscle borders should be given weights equal to 1 and 0 to landmarks corresponding to missing edges.

One can further cite [Chui and Rangarajan, 2000] a Robust Point Matching method (RPM) on its own, where the authors present a robust algorithm which not only tolerates noises but also attests of a certain ability in overcoming local minima and bad initializations. There is also the ASM version of the latter paper [Abi-Nahed *et al.*, 2006] that combines point distribution model both in 2D and 3D with robust point matching while performing image global search for feature points of interest.

The vast majority of existing approaches are explicit/landmark-based, where some critical points are deducted along with an interpolation strategy towards describing the shape [Bookstein, 1997c]. Examples refer to active shape models [Cootes *et al.*, 1995], spline-driven representations [Bookstein, 1989], triangulation of surfaces [Lamecker *et al.*, 2004], wavelet-based representations [Nain *et al.*, 2007], among others. An example for accurate segmentation achieved through deformable models combined with shape modeling is shown in [Cremers and Rousson, 2007].

Active Shape Model (*ASM*) represent the gold standard of the statistical shape models (*SSM*) with a key number of publications [Cootes *et al.*, 1995] [Cootes *et al.*, 1994a] [Cootes *et al.*, 1992] [Cootes *et al.*, 1998a]. They are mainly based on a Point Distribution Model (*PDM*), and used to both capture and represent the shape variation in training set. Active Shape Model (ASM) [Cootes *et al.*, 1998a], build a point distribution model, which allows for global scale analysis of shape variation by applying principal component analysis (PCA) to the positions of the boundary points, while using local appearance to fit the model to new data. There is also lot of work on shape priors, including Fourier representations of contours [Staib and Duncan, 1992], atlas-based segmentation methods, level set and priors [Cremers and Rousson, 2007].

The landmarks have to undergo a process of manual annotation, followed by an alignment into a common coordinate system. Principal Component Analysis is then applied in order to reduce the dimensionality of the data set and to specify the coordinate frame of

orthogonal principal axis. Consequently the shape model allows for considerable meaningful variability. One of the most significant contribution of *ASM* is that the model is able to learn the characteristic pattern of the shape class.

However, there exist some limitations related to this approach. First of all, their inability (with a relatively few number of eigenvectors) to capture the full range of anatomical shape variability. Moreover, one has to refer to their sensitivity to partial occlusions and inability to deal with orientation changes. A reduced number of examples in the data set is also considered as a handicap. In fact it is a hard to estimate a high-dimensional probability distribution of a shape starting from a relatively small number of samples. More simply, the number of eigenvectors necessary for shape representation has to be at most equal to the number of training examples minus one, while the number of points comprising the shape might be two or more orders of magnitude higher.

These limitations are reported in several reviews in the literature such as [Davatzikos *et al.*, 2003] who propose to overcome to the first drawback by calling upon a hierarchical concept based on a wavelet transform. This new formulation of the ASM is therefore capable of capturing fine and coarse variations. Duta and Sonka also present another improvement of ASM by integrating an *a priori* knowledge concerning the neuro anatomical structures of the brain to resolve these problems in [Duta and Sonka, 1997]. In [Cootes *et al.*, 1994b] a multi resolution implementation is developed to improve ASM code in terms of robustness, accuracy as well as speed. There is consequently a continuous demand on more accurate and automatic approaches of shape modeling.

To summarize this global framework, such a process involves three key aspects: (i) Shape representation, (ii) Modeling shape variations, and finally (iii) Inference.

Once the representation has been been defined, the next step consists of learning a manifold on this space from a set of training examples. Linear subspaces, parametric as well as non-parametric densities have been considered to model shape variation through the observed global distribution of the landmarks within the training examples. Single and multi-variate Gaussians, and kernel-based representations of fixed and variable bandwidth are some examples.

During the search, the inference step consists of recovering an instance of the representation being part of the model manifold that is best supported from the observed image features.

Often, these three steps are treated independently. Once the representation has been

determined, an assumption on the statistical model is made and the parameters of the manifold are determined from the training set. This manifold is then used along with the image features for object extraction. One should therefore make the following observations:

(i) a strong dependency exists between the representation and the model since by changing the distribution of control points, the model can be very different with regard to its representative capabilities,

(ii) a strong dependency exists between the representation and the segmentation since image features are often computed along this representation,

(iii) a strong dependency exists between the model and the segmentation .

For a more exhaustive review of 3D statistical shape models for medical image segmentation we strongly recommend the recent review [Heimann and Meinzer, 2009]. For a detailed version about ASM and AAM we refer the reader to [Cootes and Taylor, 2001a].

3.3 Shape Model Construction

3.3.1 Shape Representation

As far as medical image analysis is concerned, a large variety of shape features were proposed throughout the literature. One can consider those features in different categories, like dense surface meshes [Kelemen *et al.*, 1999], deformation fields [McInerney and Terzopoulos, 1996], Fourier surfaces [Staib and Duncan, 1996] as extension of 1D/2D Fourier transforms, or landmarks [Bookstein, 1997a].

Throughout most parts of this work we use point distribution models (PDM), and rely on landmarks as the representation of shapes in the image data. Landmarks are typically points that are either labeled manually or extracted automatically from the image data. Landmarks are points that have corresponding positions across a set of images (e.g., a landmark is always located on one specific anatomical location). They are a prerequisite for various landmark-based image registration and appearance based models.

While initial work relied on manually place landmarks during the training of shape models [Cootes *et al.*, 1992], subsequent approaches place landmarks in a semi-automatic [Davies *et al.*, 2002a] or fully automatic manner in the training data [Cootes *et al.*, 2005] [Langs *et al.*, 2007].

The choice of landmarks - or of subsets of landmarks used during search - should satisfy several properties: 1. Given the landmarks it should be possible to reconstruct the entire structure or object with maximal accuracy, i.e. they should cover the object sufficiently densely, and 2. The landmarks should correspond to image content that allows for un-ambiguous localization. The selection of the landmarks as well as the interpolation strategy is an important challenge towards recovering the most compact complex representation with the best possible geometric reconstruction of the object under consideration.

Suppose we are given a set of training volumes or images

$$\mathbf{I}_1, \mathbf{I}_1, \ldots, \mathbf{I}_{N_T}. \tag{3.1}$$

For a set of landmarks positions are known in all examples. That is, for each example $\mathbf{I}_i$, the landmarks are located at the positions

$$\mathbf{X}_i = \{\mathbf{x}_1^i, \mathbf{x}_2^i, \ldots \mathbf{x}_{N_L}^i\}. \tag{3.2}$$

where $\mathbf{x}_i \in \mathbb{R}^d$. We will call $\mathbf{X}_i \in \mathbb{R}^{d \times N_L}$ a shape, and denote the set of shapes in the training set by

$$\mathcal{X} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{N_T}\}. \tag{3.3}$$

3.3.2 Alignment and Procrustes Analysis

A prerequisite to any shape representation analysis is to align the data set into a common coordinate frame [Goodall, 1991]. This normalization allows for getting rid of the differences across shapes that are mainly due to rotation and translation. Procrustes analysis [Gower, 1975] represents the gold standard of the alignment methods in the literature. More practically, it minimizes the Procrustes distance, representing a least square measure of the shape dissimilarity, between geometric shapes. There is also an effective and

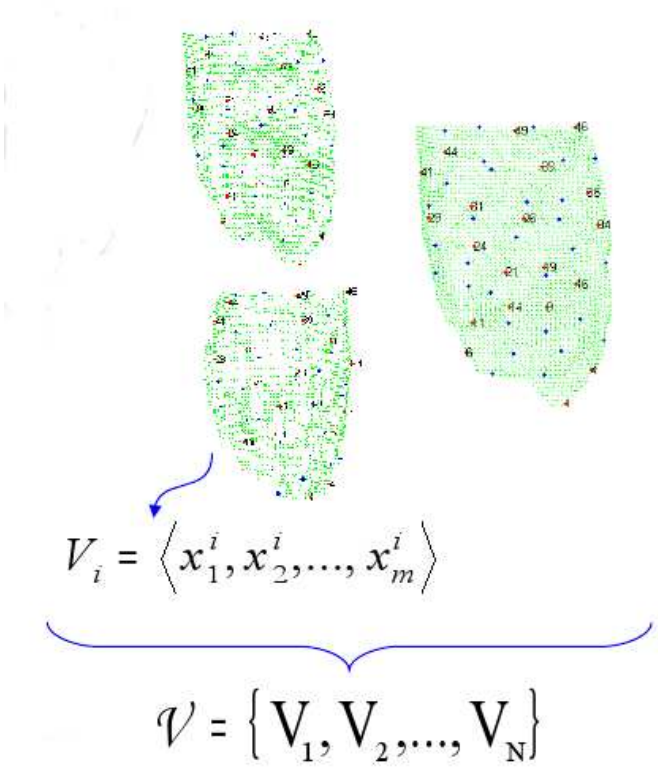


Figure 3.2 — Point Distribution Model

simple iterative approach [Cootes and Taylor, 2001a], whose steps can be summarized in Alg. 1.

Some downsides of the Procrustes analysis remain yet unresolved, especially the problem of missing landmarks. For a relevant study on alignment and some answered drawbacks see [Ericsson, 2006].

3.3.3 Obtaining Correspondence

3.3.3.1 Overview

An accurate correspondence of the training volumes is a pre-requisite for a rigourous point-based shape model. In fact a dense correspondence is even a key step in the model building process. Landmarks are in practise said to correspond if they are matching in

Algorithm 1 Iterative Procrustes Analysis Algorithm

1. Translate the training set examples so that their centers of mass are at the origin.
2. Opt for one example as initial guess of the mean shape, $\bar{x}$, and scale it so that $\|\bar{x}\| = 1$.
3. Record the first estimate as $\bar{x}_0$ to define the default reference frame.
4. ***Repeat***
5. Align all the shapes regarding to the estimated mean shape.
6. Estimate a new mean shape $\bar{x}$.
7. Impose the new mean shape by aligning it with x_0 and scale it so $\|\bar{x}\| = 1$.
8. ***Until Convergence***

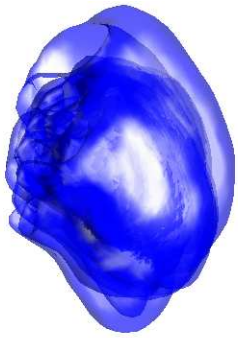
a biological meaningful way. If all landmarks are located on a manifold i.e., the surface of a muscle, we imply by corresponding positions that the following landmarks;

$$\mathbf{x}_j^1, \mathbf{x}_j^2, \ldots, \mathbf{x}_j^{N_T} \tag{3.4}$$

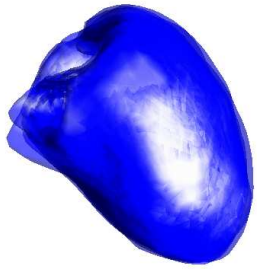
have to be located at the same anatomical position, in all training examples. This correspondence is established during training by either taking only the shape into account, or by using both shape constraint of a continuous annotation of the shape (e.g., anatomical object) and the image content in its vicinity.

As far as the 2D correspondence problem is concerned, manually established landmarks can be determined [Cootes and Taylor, 1995] [Bookstein, 1997b] which certainly leads to reasonable and satisfactory results. However it seems to be a time consuming task and even a source of multiple errors, leading to a poor model generation.

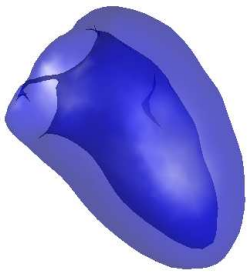
As explained in [Davies, 2002], the straightforward approach to define correspondence is to select a starting point on each example and equally space a number of points on each boundary. Actually different conceptions could be derived in order to solve this problem of correspondence, it could be distance-based [Golland *et al.*, 2000], shape (boundaries)-based [Hill *et al.*, 2000], or still image-based [Rueckert *et al.*, 2003].



(a) Training set before alignment.



(b) Aligned left ventricle.



(c) Heart Mean Shape.

Figure 3.3 — Left Ventricle Alignment through Procrustes Analysis.

For a complete comparison and explanation of these correspondences methods, we refer the reader to [Davies, 2002], however it is noteworthy that the distance-based techniques have an important side effect in a way that the relative position of equivalent points can change noticeably over the data set. Regarding the non-rigid image registration, they lead to reasonable but arbitrary results. This is essentially due to the existence of an infinite number of non-rigid deformations adequate to match the intensities in two images.

As for 3D correspondences, they can be for example obtained by semi-automatic Minimum Distance Length -*MDL*- proposed in [Davies *et al.*, 2002a], fully automatic technique as in [Cootes *et al.*, 2005] [Langs *et al.*, 2007], the congealing method [Zollei *et al.*, 2005], or the covariance determinant (DetCov) [Kotcheff and Taylor, 1998], and finally spherical harmonics technique (SPHARM) [Brechbuehler *et al.*, 1995].

The DetCov method possess the specificity of minimizing the covariance matrix, and

therefore gives privilege to compact models, but presents the inconvenient to have an exorbitant computational cost. Whereas the SPHARM is especially dedicated to objects with spherical topology, by having the spherical harmonics as the basis functions of the parameterized surface. This approach was used in [Kelemen *et al.*, 1999] where the authors propose to build shape models starting from a set of closed 3D shapes by defining correspondence through every surface parameterization. Another recent technique presented in [Chui and Rangarajan, 2000] deals with the correspondence issue, however -yet like many algorithms- it only matchs one volume to another, rather than considering the whole data set. Typically this kind of method do not guarantee continuity neither bijectivity. Hence they are not really effective for shape modelling. In other words the parameterisation functions do not have to be monotone and not all points need to be matched.

A thorough survey of 3D methods for landmark correspondence can be found in [Styner *et al.*, 2003], where the latter three methods are evaluated regarding several criteria: compactness, generalization and specificity. Among the outcome of this evaluation we point out the supremacy of the Minimum Distance Length -*MDL*- based approach representing a method for inductive inference. The details of the latter method will be therefore discussed in the upcoming section.

3.3.3.2 Minimum Distance Length

The Minimum Description Length (MDL) [Davies *et al.*, 2002a] approach scores well the issue of points correspondence. The MDL method seeks to achieve a compact description of the landmarks positions in the shape. In case we want to establish how well our data set can be modeled by a specific model family, we have to first assume a standard multi variate Gaussian model (such in ASM or AAM). The basic idea would be to have a model which is able to compress data, then we assume that it captures or approximates the model underlying the data. Thus a higher compression ratio leads to a better fitting of the data and the model.

Practically the MDL is originally based on the DetCov technique [Kotcheff and Taylor, 1998] in the way that it was introduced through a different objective function for the optimization process.

The principle of MDL is based on transmitting a set of shapes as an encoded message, where the code emanates from a prefixed set of parametric statistical models. Afterward

the transmission incorporates the encoded data values as well as the coded model parameters.

Hence MDL equilibrates the model complexity, which is explicitly formulated thanks to the transmitting cost of the model parameters, against the fitting quality between both the model and the data, determined by the coding length. More simply the MDL represents the necessary effort to send the model bit by bit as shown in Equation. (3.5);

$$L(D, \mathcal{M}) = L(\mathcal{M}) + L(D|\mathcal{M}) \tag{3.5}$$

where $L(\mathcal{M})$ is the cost of communicating the shape model and $L(D|\mathcal{M})$ is the cost of the shape data encoded with the help of the model.

3.3.3.3 Robust Autonomous Model Learning

We have reviewed during the last section the major previous work on correspondence establishement between shapes, especially those handling the population based optimization issue. Although these methods produce plausible as well as partly reliable results, the major remaining contraint is relative to the choice of topology and parameterization. The Minimum Distance Length technique [Davies *et al.*, 2002b] for instance utilizes an a priori chosen topology to parameterize correspondences and deformation throughout the learning and the search process.

As our data set, e.g. calf muscle, presents a high variability of local deformation behavior, we called for a more weakly supervised learning algorithm for appearance models [Langs *et al.*, 2007] based on the minimum description length (MDL) principle, to resolve the correspondance problem.

Among the benefits of this method, one can cite that we have;

- No need for annotation (a time consuming and error prone task).
- No hypothesis about the topology of the shapes,
- No continuous representation of the volumes, during optimization the method only requires a distinctive sets of interest points.

The following automatic model building is even adequate to deal with missing landmarks issue, addressed by means of a robust model estimation. Indeed the idea behind

this technique is based on a group-wise registration of sparse representations of the training data.

Pre-Processing: Group-Wise Registration Here is the outline of the algorithm developed by [Langs *et al.*, 2007]

1. Pairwise matching of I_1 to $N_T - 1$ images with K interest points/landmarks.
2. Correspondance outcome is stored in $\mathbf{G} \in \mathbb{R}^{k \times n}$.
3. Starting from this correspondance result, the authors proceed to a group wise registration that minimizes a criterion function. The latter criterion is based on the compactness of the appearance model, depicting both shape and local texture variation.

In fact the criterion function is inspired from the Minimum Distance Length idea, however instead of using a two-part coding scheme, as in Equation. (3.5), the description length is decomposed into multiple parts, as in Equation. (3.6);

$$\mathcal{C} = \mathcal{C}_S + \mathcal{C}_T + \alpha(t)\mathcal{C}_E \tag{3.6}$$

including a shape compactness criteria $\mathcal{C}_S$ (Equation. (3.7)), a local texture parameter $\mathcal{C}_T$ (Equation. (3.8)), and finally an elasticity regularization measure $\mathcal{C}_E$ (Equation. (3.9)), as detailed in the following equations:

$$\mathcal{C}_S = L(\mathcal{M}_S) + L(D_S|\mathcal{M}_S) + \mathcal{R}_S - Lref \tag{3.7}$$

$$\mathcal{C}_T = L(\mathcal{M}_T) + L(D_T|\mathcal{M}_T) + \mathcal{R}_T \tag{3.8}$$

$$C_E = |\nabla \mathbf{d}(\mathbf{x})|^2 \tag{3.9}$$

where :

- $L(\mathcal{M}_S)$ and $L(\mathcal{M}_T)$ represent the transfer cost of the shape model,

- $L(D_S|\mathcal{M}_S)$ and $L(D_T|\mathcal{M}_T)$ consider the cost of the shape data encoded with help of the model,
- $\mathcal{R}_S$ and $\mathcal{R}_T$ as a penalty for the residual error not captured by the model.
- $L_{ref} = \sum_{i=1,\ldots,N} entropy_{j=1,\ldots,k}(\mathbf{p}_{ij})$ helps for the normalization and constitutes the entropy of the landmark positions $p_{i,j}$ in the each volume.
- d the displacement of any landmark throughout the data set, that way we prevent a degenerate model.

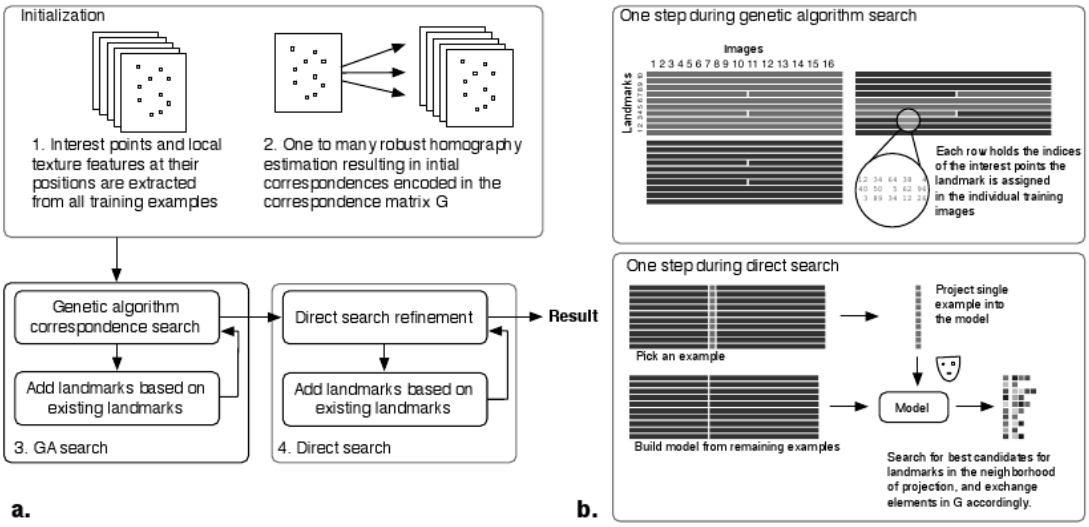


Figure 3.4 — Description of the Robust Autonomous Model [Langs *et al.*, 2007]

Afterward the optimization is insured through, at first, a genetic algorithm search coupled later on with a fine search, and starting with a limited number of landmarks, while adding some others just after the learning process, see Figure. 3.4 for more details.

3.3.4 Dimensionality Reduction

Whenever we face a huge dat set, there is in reality only a small number of parameters that govern the data, which constitute the true dimension of the training set as well as a real motivation for the data reduction. In the most typical cases we call upon the dimensionality reduction techniques that were studied extensively in the literature including

PCA (Principle Component Analysis) or rather ICA (Independent Component Analysis) [Hyvärinen and Oja, 2000]. The PCA represents though the most common technique to decompose the shape space and that maximizes the variance of the input data. The combination of PCA and feature vector has proved to work [Cootes *et al.*, 1994a] in terms of anatomical segmentation.

In practice each step during the PCA performs a variance maximizing rotation of the original space. It generates in this manner new ordered axes according to the variance of the shapes in the different axis directions as seen in Figure. 3.5. That way the principal components exhibit a natural ordering according to the variance they describe. This kind of model is able to capture effectively considerable variability.

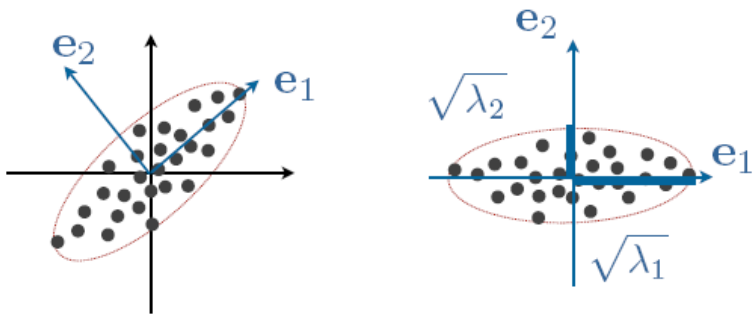


Figure 3.5 — PCA eigenvalue decomposition

Nonetheless there are some constraints with regard to the PCA approach, such as the size of the training set. The latter has to be consistent in order to capture the variation within a same class. Besides the assumption of the Gaussian probability distribution over the data generation has to be assessed. Moreover, PCA works to encode global variations only, whereas more local variations are not captured. This is directly related to the fact that the eigenvectors of the covariance matrix the most relevant modes of variation.

Before starting to reduce the dimensionality of the data set, we first represent each data by its deviation from the mean $\bar{\mathbf{x}}$

$$\bar{\mathbf{x}} = \frac{1}{N_T} \sum_{i=1}^{N_T} \mathbf{x}_i \tag{3.10}$$

We perform then a Singular Value Decomposition (*SVD*) over the covariance matrix

S -see Equation. (3.11)- of the aligned data in order to evaluate the principal components of the training set. Let's suppose we have aligned the shapes to $\mathbf{x}_1$, $UDV = \mathbf{x}_i^T \mathbf{x}_1$ where VU^T is actually the rotation matrix.

$$\mathbf{S} = \frac{1}{N_T - 1} \sum_{i=1}^{N_T} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \tag{3.11}$$

Once we capture the non-rigid shape variation, global translation, scale and rotation do not affect the model complexity. The aligned shapes are modeled by a multivariate Gaussian with model mean $\bar{\mathbf{x}}$ and eigenvectors of the covariance matrix

$$\mathbf{x}_i = \bar{\mathbf{x}} + \sum_{j=1}^{N_P} b_j^i \mathbf{e}_j \tag{3.12}$$

where:

- $\mathbf{e}_j$ represent the modes of variation of the covariance matrix $\mathbf{S}$.
- b_j^i the coefficients symbolizing each shape $\mathbf{x}_i$
- N_p corresponds to the number of the largest eigenvalues, generally this number is chosen in a way that a certain ratio of 98% of the variance is shown in the training data set. While varying the modes we can establish a given limits $\pm 3\sqrt{\lambda_i}$ [Cootes and Taylor, 2001a] as shown in Figure. 3.6, with λ_i the corresponding eigenvalues.

An alternative way to constrain the shape parameters to a determined interval is shown in Equation. (3.13)

$$\left(\sum_{i=1}^{N_P} \frac{b_i^2}{\lambda}_i \right) \leq M_t \tag{3.13}$$

where M_t stands for a threshold taken from the χ^2 distribution. That way the group of modes are considered as one multivariate distribution, and the shape parameter will be bound to stand inside a hyperellipsoid.

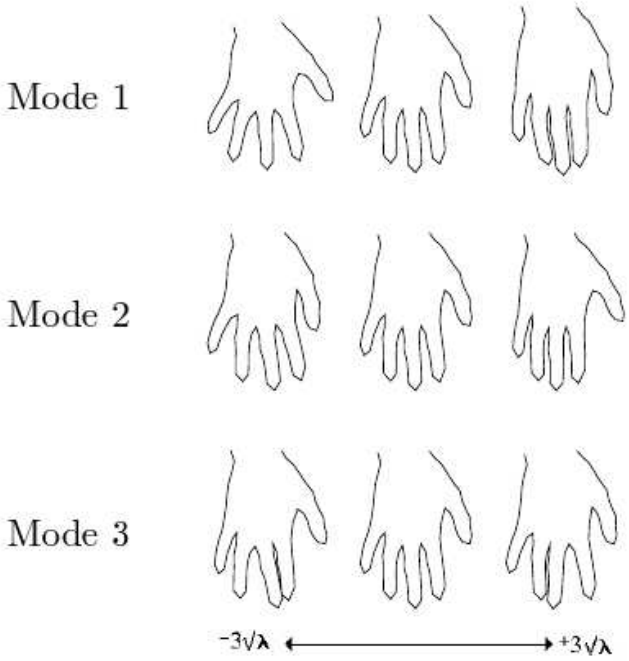


Figure 3.6 — Shape modeling example with typical hand training set [Cootes and Taylor, 2001a]

3.4 Shape Model Search

The initialization issue is always requiered for the shape models local search algorithms. The manual interaction remains a solution that can be applied as for [Kelemen *et al.*, 1999] and [Pizer *et al.*, 2003]. The construction of an atlas through an affine registration constitutes a different option as presented in the knee shape model segmentation [Fripp *et al.*, 2007] .

To answer the question how to fit a model to new points, typically we will search to define an instance of the model $\mathbf{y}$ in an image through the use of a similarity transformation T including a rotation, a translation and a scaling, associated with the shape parameters $\mathbf{b}$ as well as the eigenvectors (Equation. (3.14)).

$$\mathbf{y} = T(\bar{\mathbf{x}} + \mathbf{eb}) \tag{3.14}$$

The main goal now would be to find the optimal couple of pose and shape parameters,

which is achieved thanks to the following algorithm (Algorithm. 2):

Algorithm 2 Active Shape Model Search Algorithm

1. Initialise shape parameter $\mathbf{b}$ to zero.
2. Produce the model $\mathbf{x}$ as $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{eb}$.
3. Estimate appearance model along the normal vector to the surface at different landmarks positions to find the best similarity transformation T.
4. Update shape model parameter in a way that $\mathbf{b} = \mathbf{e}^T(\mathbf{y} - \mathbf{x})$
5. Impose constraints on shape parameter as explained in Section.3.3.4
6. ***Repeat Until Convergence***

Different variants and extensions of the ASM search algorithm have been proposed in the literature, such as the coarse to fine strategy [Cootes *et al.*, 1994b] where the starting point of the search process is the coarsest level, as soon as a convergence is met, the algorithm moves on the following level. This extension appears to be fatser and more powerful for fitting a new model than the classic Algorithm. 2.

One of the issues inherited by the active shape model search is dealing with outliers and preserving stability. For instance, either by focalizing on how to reduce outliers effects [Rogers and Graham, 2002], or by establishing a robust point matching algorithm that excludes outliers [Abi-Nahed *et al.*, 2006] to set the optimum fitting model.

Apart from the local search, there has been work on global search launched by [Hill *et al.*, 1992] who get inspired from the genetic algorithms. However one has to keep in mind that global search is a constraining task seeing that the space of search is large and local minima are multiple.

Concerning generative models, and more precisely the Active Appearance Model search (AAM search), we will handle it more into details in Section.6.2, where the chapter is dedicated to appearance based search.

3.5 Diffusion Maps

Diffusion maps are closely related to Markvo processes. They capture the structure of a graph by means of a diffusion distance [Coifman and Lafon, 2006], and embed the graph nodes into a space whose geometry reflects the mutual diffusion distances between nodes. Thereby they offer a way of dimensionality reduction. A Markov chain is constructed via pairwise similarities,which are defined locally. A diffusion process on the graph defines the geometry determined by the graph edge weights. Similar work ranges from Laplacian eigenmaps [Belkin and Niyogi, 2003], another reduction algorithm that is highly related to diffusion maps [Schclar, 2008], or diffusion kernel work on graphs [Kondor and Lafferty, 2002]. Among the previous applications related to diffusion maps one can cite text classification and lip reading [Lafon and Lee, 2006]. In the following, we will explain the basics of diffusion maps. They will be used in Section. 4.4 and Section. 5.4 to capture the structure of landmark behavior.

Random Walk Let's consider a set of point $\Gamma = \{x_k\}_{k=1}^{N}$ with $x_k \in^{n}$. We aim at walking randomly through this data set while taking an arbitrary starting point. We denote by p_{ij} the probability of transition between two consecutive points/landmarks i and j in a period of time t. One of the similarity type that can be taken into consideration is the space closeness, in other words it will be the inverse of the Euclidean distance. In such a case the probability of walking in the same cluster of points is bound to be higher than the one of traveling from one cluster to another.

Markov Process The random walk described here in represents indeed a special case of Markov process. The transition probabilities will be stored in a matrix of dimension $N.N$ where $\mathbf{P} = p_{ij}$.

For the need of diffusion maps algorithm, we will model the Markov on a graph $G = (V, E)$, where the nodes correspond to the landmarks and the edge weights to the transition probabilities. Naturally, the more similar a pair of landmarks is, the higher the edge weight will be.

The construction depicted in the Algorithm. 3 captures local geometry as well as interesting geometric features. Furthermore as the random walking is running, as the local geometry information is propagated [Coifman and Lafon, 2006] [Schclar, 2008]. The final step of eigen decomposition enables to designate a natural embedding of the data

Algorithm 3 Diffusion Maps Algorithm

Input: Data set Γ,
Pairwise similarities.

1. Building of an undirected graph $\mathbf{G}, \mathbf{W}$ on the data set Γ with the weights corresponding to the pair wise similarities between the landmarks or the different states.

2. Processing of a random walk through the graph $\mathbf{G}$ leading to a markov transition matrix $\mathbf{P}$.

3. Eigen decomposition of the markov chain matrix $\mathbf{P}$.

Output: Eigenspace (modes, eigenvalues) of the markov chain matrix.

through the diffusion map, and the eigenvectors will allow afterward for a dimensionality reduction process and hence an embedding in a lower dimension space.

3.6 Examples of Muscle Segmentation

3.6.1 Prior Art in Calf Muscle Segmentation

Muscle-compounds (see Figure. 2.5) present a rather different and new challenge to segmentation algorithms because there is no prominent difference of tissue-properties between neighboring muscles. Border tissues in between muscles are only visible on specific locations, distributed in a very sparse and heterogeneous manner. Indeed muscle surfaces are only partially visible, while parts exhibit a structure that can change dramatically between patients or during the course of follow-up examinations. Although dominated by the global anatomy, muscle deformation exhibits mostly locally consistent behavior, precluding the use of e.g., a global linear model.

State of the art medical segmentation methods mainly rely on a clearly defined topology, and an object boundary characterized by salient features (e.g, edges) [Duncan and Ayache, 2000]. The musculoskeletal modeling problem in medical imaging is not widely investigated in the literature. Indeed few works have been dedicated to this issue [Gilles *et al.*, 2006] [Blemker *et al.*, 2007]. Related work on muscle segmentation includes simplex meshes [Gilles *et al.*, 2006], in fact a mesh in computer graphics represents a

collection of vertex points sampling a hyper-surface, and connected by edges, faces or hyper surfaces. Another method presented in [Fernandez and Hunter, 2005] explores an atlas-based models which are deformed using a free-form deformation . The main challenges of calf T1-MRI segmentation is the total absence of conventional image support. Edges between classes are almost invisible while at the same time the global intensity characteristics of the muscle regions are almost identical. In fact the border tissues in between muscles are only visible on specific locations, distributed in a very sparse and heterogeneous manner. Indeed muscle partially exhibit structures that can change dramatically between patients, or during the course of follow-up examinations. Nevertheless, and as highlighted by [Blemker *et al.*, 2007], musculoskeletal disabilities in general and Myopathies -as far as our work is concerned- could highly profit from this kind of studies and improve future treatments.

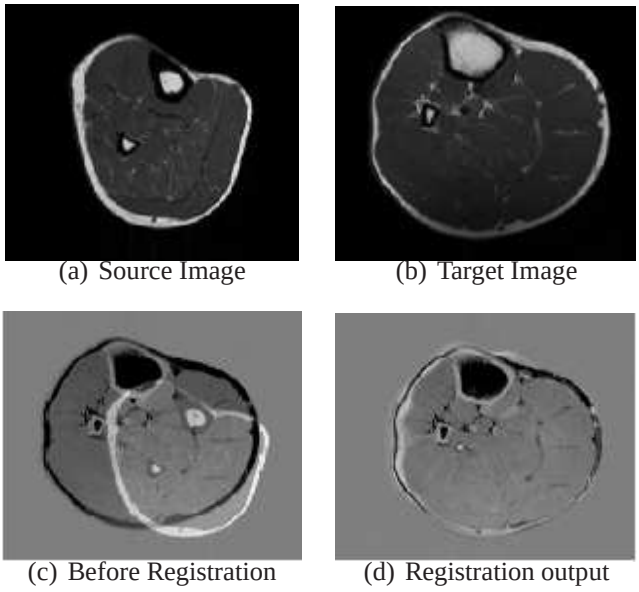
Consequently prior models build from conventional image-based search approaches will fail to separate the muscles regions. An alternative would be the explicit use of the image support during the construction of the model. In fact selective active shape models for example [Cootes *et al.*, 1994a] use only a small portion of the landmarks during the search, while sparse models [Florin *et al.*, 2007] aim to optimize shape representation by proposing a sparse representation that encodes sparsity and exploits geometric redundancy choosing individual representative slices taking into account image support. These methods depend heavily on the accuracy of the inter-subject registration for group comparison and the parametrization of the shape.

Registration-based Segmentation Initialization We perform a uni-modal registration of an annotated calf MRI (the atlas) to the target volume through the use of a novel deformable registration framework [Glocker *et al.*, 2008]. The registration problem is formulated as a markov random field (MRF) optimization and the approach is based on discrete labelling and linear programming, which presents the advantage to be gradient-free and flexible in the choice of the distance measure. This technique also guarantees optimal properties on the solution, computational efficiency and tractability. Assuming a prior model that involves both geometry and texture (segmented anatomical atlas), one can define the segmentation through the deformation of the model to the image that is a natural registration problem, and can be addressed by the proposed framework.

A scheme of registration based segmentation of muscles in MRI data is depicted in Figure. 3.7. The expert annotation of an example is transfered to a target by non-rigid

registration.

The registration obtains a coarse initial segmentation and results indicate the good behavior of this method on our data set, while at the same time the need for a more accurate segmentation that can not be achieved by the registration. Although the technique achieves a good appearance matching as shown in Figure. 3.8, the segmentation result is still not rigorous enough, and there is still a need for more veracity and precision.



(a) Source Image (b) Target Image

(c) Before Registration (d) Registration output

Figure 3.7 — Registration result for MRI calf, the source and the target images in the upper row and the difference visualisation before and after the registration in the lower row.

3.6.2 Prior Art in Left Ventricle Segmentation

The automatic delineation of the Left Ventricle (LV) is a critical component of computer-assisted cardiac diagnosis. Information with respect to the ejection fraction, the wall motion and the valve behavior can be very useful toward predicting and avoiding myocardial infarction as shown respectively in [Kaus *et al.*, 2004] [Jolly *et al.*, 2001]. During our study we will mainly focus on CT images of the left ventricle. Concerning the Computer Tomography, the main applications of this modality in cardiac imaging are the evaluation of cardiac masses as well as the evaluation of aortic and pericardial diseases.

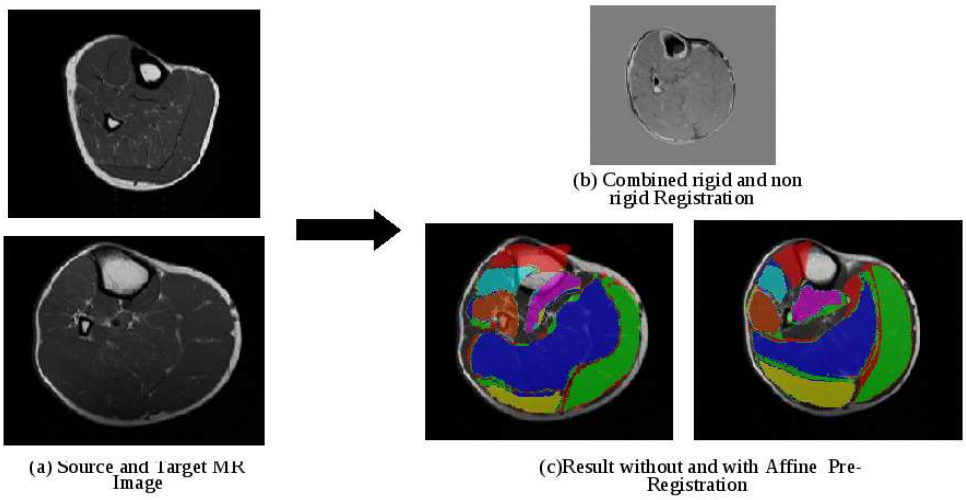
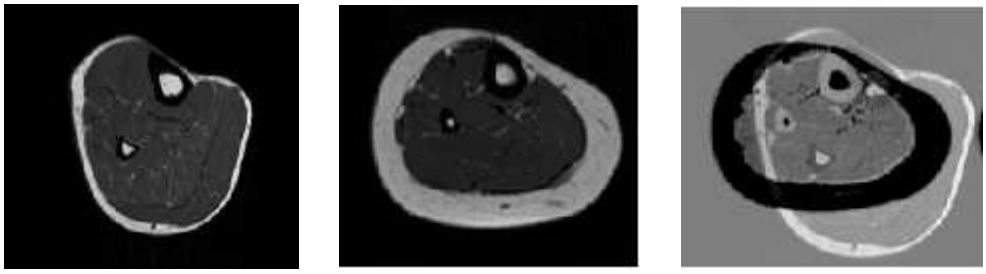
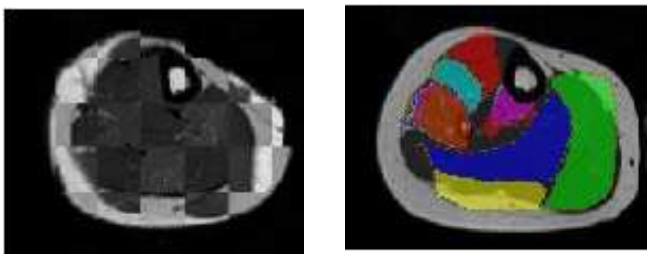


Figure 3.8 — Segmentation-based Registration result on calf muscle



(a) Source Image (b) Target Image (c) Before Registration



(d) Checkerboard visualization (e) Segmentation based registration result

Figure 3.9 — Another example of segmentation based registration result with no need for affine pre-registration.

The segmentation of the left ventricle is challenging mostly due to the similar visual properties with the other chambers of the heart cavity, as well as the presence of papillary muscles. In this context the use of edge-driven terms with regional statistics along either with deformable contours or active shape and appearance models would not lead to successful results. In the first case, computational complexity is an issue and the proper handling of papillary muscles is problematic. In the second case, one has to deal with either the linearity of the sub-space or the fact that building appearance modes requires appearance normalization and too many samples.

Model-free and model-based methods were widely investigated in the past, and we refer to [Cremers *et al.*, 2002] for a rather comprehensive survey. The use of the shortest path algorithm along with shape matching was considered in [Jolly, 2006], while Staib et al. have introduced a 3D ballon model in [Staib and Duncan, 1996]. The latter model is parameterized on an orthonormal Fourier basis in a way to allow for a constrained image search. Thus the model fitting is achieved by balancing an internal energy term with an external gradient-derived scalar field. Another effective method to model anatomic data is described by the M-Reps approach (medial model representation) [Pizer *et al.*, 2003] in a hierarchical coarse to fine fashion.

Statistical point distribution models were also devoted to solve cardiac segmentation issue by Cootes and Taylor in [Cootes *et al.*, 1992], as well as in [Cootes *et al.*, 1995], combining in this manner shape and boundary gray-level aspects. As the image appearance information was still not exploited, despite its importance and richness, it was finally introduced in Active Appearance Model (*AAM*) [Cootes *et al.*, 1998b], and among applications figures the spatio-temporal heart segmentation for the 2D case [Mitchell *et al.*, 2001] as well as the 3D case in [Mitchell *et al.*, 2002]. Last, but not least, we refer to [Paragios, 2002] for an alternative shape representation using level set functions.

For a complete review, we refer the reader to [Jolly, 2006], [Frangi *et al.*, 2001] and references therein.

Nevertheless one has to bear in mind that in all the above-referenced approaches, an a priori topology is required in order to build a segmentation model.

3.7 Summary

In this chapter, we provided an overview of the background and the state of the art of the medical segmentation problem.

Furthermore, we highlighted the methods commonly used in the case of the calf muscle and the left ventricle segmentation. Despite all the pros and cons of the segmentation approaches described previously, they generally suffer from the lack of compatibility with our image data, especially that the image support quality is not really taken into account. In the vast majority of segmentation techniques, it is presumed that if ever an image region quality is low, another one would cerainly equilibrate.

We will point out in the next chapter our choice of a more suitable shape representation regarding the challenging properties of the studied medical images this thesis and relevant to the problem of statistical shape analysis. Therefore instead of using a priori choices, we need a model that rather learns the intrinsic shape topology from the training data, while represent the shape variations regarding to this topology.

Conventional shape modeling approaches rely on a pre-defined topology and parametrize the surface of an object with regard to an according manifold like a sphere, thus limiting themselves to specific tasks. In the ideal case, one would like a method that learns the appropriate topology from the training data and uses an according shape representation to model its potentially rather heterogeneous variation and local texture features extracted at the landmark positions. Such a powerful model has to deal with two challenges: (i) A way to determine the intrinsic topology of a shape for which multiple examples are available, (ii) Means to represent the shape variation with regard to this topology, i.e., a way to exploit arbitrary topologies.

CHAPTER

4 Sparse Shape Models

"Not everything that can be counted counts, and not everything that counts can be counted."

Albert Einstein

4.1 Introduction

The localization and segmentation of muscles in MRI data poses several challenges. The lack of contrast between the anatomical regions as well as the highly heterogeneous distribution of salient features makes the segmentation task almost intractable with conventional methods. In the following we will explain how to to exploit redundancies in the shape variation and use the image information efficiently to obtain reliable segmentation results. In this chapter we propose a sparse representation that encodes sparsity, exploits geometric redundancy, and most importantly accounts for the varying degrees of image support to obtain a robust segmentation of individual anatomical structures.

Indeed a novel technique for model-based vision is introduced that performs dimensionality reduction while taking the image support into account. The fundamental assumption of our method is that one can describe anatomical structures while just referring to a small number of elements, and reconstruct the entire structure through a well chosen, and object specific through interpolation. Hence we have to find the smallest possible set of robust, most representative, best supported components and features capable to retrieve an optimal reconstruction of the original object through a data-driven interpolation method. The resulting segmentation serves as a basis for the consistent localization of regions of interest in the anatomy. This is relevant in cross-sectional studies, and during follow-up examinations.

Throughout this section one of your goals is to answer some critical questions:

1. How can we learn the structure of behavior and how can we use this in shape and appearance models ,
2. How can we measure the local image information that supports segmentation,
3. How can we integrate this knowledge about the structure and redundancies in the shape variability, and the distribution of image support to achieve an optimal sparse representation of the objects.

The remainder of the chapter is organized as follows: in Sec. 4.2 we recall some related work regarding the sparse techniques for building more localized shape models, whereas in Sec. 4.3 the optimal shape representation and model construction based on the data sparsity are presented. Afterward Sec. 4.6.2 deals with the inference and the optimal use of image support.

4.2 Related Work

In the same line of research we find some interesting work in the literature;

- An alternative parametrization of shape is Independent Component Analysis [Üzümcü *et al.*, 2003] was investigated for cardiac MR images segmentation and also compared to the principal components analysis results. Among the advantages of ICA decomposition one can cite that it does not require a Gaussian distribution of the input data and that is capable to capture localized shape variations.

- [Florin *et al.*, 2007] focused on liver segmentation and called for the level-set techniques for 2D key slice segmentation and then a 3D shape was interpolated from 2D contours. They introduced 3D sparse models and distinguished between two classes of regions present in the object; low support regions, and high support ones. In fact they even propose to extend the notion of image support toward being associated with the quality of the samples in general. The main purpose remain to select robust regions with high support, and to be able to reconstruct the object of interest just with these extracted elements. However the approach is not point distribution based, it rather accounts over key number of slices extracted from a 3D volume.

- The SPCA decomposition standing for Sparse Principal Component Analysis [Sjostrand *et al.*, 2007] extracts sparse and meaningful anatomical elements from a training data set. In fact one of the inconvenient of the principal component analysis is that the principal component contained in the loading matrix are non-zero, which makes the interpretation arduous. The scope of SPCA is so far to make each PC to be dependent on a limited set of variables [Sjöstrand, 2007]. The approach can be conceived as regression type optimization problem of the PCA using LASSO, a selection variable technique leading to sparse models [Zou *et al.*, 2006].

- Another method for building more localized shape models was introduced in [Leung and Bosch, 2007b] for the classification of local wall motion abnormalities in left ventricle; a comparison between different orthomax criterion (varimax, quatrimax, factor-parsimony) was carried. The method relies on the rotated shape space of the PCA, where could be found localized spatial variations. For more details

concerning the orthomax rotation we refer the reader to the upcoming chapter and more precisely to the Section.5.5.1.

There has also been work on statistical shape analysis using non-Euclidean metrics by [Larsen and Hilger, 2003] to obtain sparse modes of variation, where the authors even proceed to an extensive comparison between PCA, Maximum Autocorrelation Factor (MAF), and Minimum Noise Fraction (MNF).

4.3 Sparse Shape Model Concept

Sparse shape models learn a representation and a corresponding reconstruction mechanism from a set of training examples. The sparse model is built based on the statistical behavior of the training shapes and the distribution of appearance information in the training data. In the following we will formulate the framework, and explain how to derive an optimal sparse representation from training examples. Subsequently, the method for the reconstruction of the entire modeled structure from the sparse representation, and will explain the search procedure.

In order to segment shapes based on a sub-set of elements (landmarks in our case), a training set of volumes or images each containing an example of the structure of interest, we define;

- A representation: landmarks for which corresponding positions are known in all examples
- A model the captures the variation of the shapes in the training set, and how to reconstruct a shape with help of this model if only a part i.e., a sub-set of landmarks is known.
- A method to select a sub-set of landmarks, that is best suited for reconstruction from new data, by taking into account the typical image support in the training examples, and the geometrical significance of individual landmarks, i.e. how important are they for reconstruction?

The landmarks allow for consistent use of particular positions on the muscle surface. With help of this representation we are able to assign individual positions on the surface

properties like expected image support, or significance for the reconstruction process, independently from a prior choice of the coordinate system.

Since there exists a correlation in the training set of shapes, we expect to be able to obtain a reconstruction of a shape within the object class, if only a part of the landmarks are known. In the following a framework for the reconstruction of the shapes, i.e., the entire set of landmarks from a sub-set is described. The choice of this sub-set of landmarks will be also explained in the following.

4.4 Sparse Shape Model Construction

Let us consider a shape representation that consists on a finite set of landmarks. Given a set of n training volumes and their corresponding segmented structure

$$\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n, \tag{4.1}$$

our knowledge about the data comprises m landmark positions in each of the examples. The number of landmarks can by high, up to a dense sampling of shape surfaces. Landmarks are not constraint to anatomically salient points, but can be distributed on manually segmented training examples by methods like those proposed in [Davies *et al.*, 2002b] [Langs *et al.*, 2007]. The landmark positions can be found using a number of approaches, such as a minimum description length based criterion. Landmarks do not have to be located on a single surface or manifold, but can define arbitrary structures and deformation fields [Taron *et al.*, 2007]. For each example $\mathbf{I}_i$, the landmarks are located at the positions

$$\mathbf{V}_i = \{\mathbf{x}_1^i, \mathbf{x}_2^i, \ldots \mathbf{x}_m^i, \}. \tag{4.2}$$

where $\mathbf{x}_i \in \mathbb{R}^d$. We call $\mathbf{V}_i \in \mathbb{R}^{d \times m}$ a shape, and denote the set of shapes in the training set by

$$\mathcal{V} = \{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n\}. \tag{4.3}$$

This data defines a shape manifold, that can be associated with geometric and image support. This should happen while satisfying two conditions:

(i) Preservation of the information necessary to reconstruct the shapes with maximal accuracy,

(ii) Image information associated to the landmarks that enables localization in new data, by having low ambiguity.

Let us define;

$$\hat{\mathcal{V}} = \{\hat{\mathbf{V}}_1, \hat{\mathbf{V}}_2, \ldots, \hat{\mathbf{V}}_n\}. \tag{4.4}$$

where $\hat{\mathbf{V}}_i \in \mathbb{R}^{d \times m'}$ are the representations of the full shapes V_i in the training set, with $m' \ll m$. $\hat{\mathbf{V}}_i$ consists of a sub set of the landmarks defining the shape, and a corresponding reconstruction function P expressed as following:

$$P : \mathbb{R}^{d \times m'} \rightarrow \mathbb{R}^{d \times m}, \hat{\mathbf{V}}_i \mapsto \mathbf{V}_i + \mathcal{R} \tag{4.5}$$

where $\mathcal{R}$ is a residual error, that should be minimal.

We will first discuss how to obtain this representation $\hat{\mathcal{V}}$ based on the shape and appearance behavior in the training set, in order to obtain optimal search ability. In this work we consider a multivariate Gaussian shape model, as used in [Davies *et al.*, 2002b], to model the shape variability.

4.4.1 Shape Maps and Redundancy

We view finding an optimal shape representation as an optimal sampling with respect to the variations being observed in the training data. It should have low density in regions that behave in a redundant manner, and high density in regions that exhibit uncorrelated or complex deformation behavior in the training set. Analogously to a uniformly distributed sampling in real space, that covers the object evenly, the sparse representation has to cover the object evenly with regard to the *information* contained in each sampling point. To achieve this, we have to capture the coherence of the behavior of shape regions in the training examples. In [Langs and Paragios, 2008] shape maps are introduced. They provide for a shape population metric, that captures the interdependencies in the behavior of landmarks. We use the concept of shape maps to derive an optimal sampling.

4.4.1.1 Markov Chain Description

Diffusion maps [Coifman and Lafon, 2006] represent a spectral embedding of a set X of n nodes, for which local geometries are defined by a kernel $k : X \times X \longrightarrow \mathbb{R}$. The kernel k has to achieve both symmetric and positivity preserving:

$$\begin{aligned} k(x,y) &= k(y,x) \\ k(x,y) &\geq 0 \end{aligned} \tag{4.6}$$

Let us consider a Markov chain consisting of m nodes, that correspond to the landmarks, and edges with a value $p_k(i,j)$ between nodes that correspond to the minimal description lengths [Rissanen, 1978] of models encompassing the two landmarks i and j and $k-2$ other landmarks. The description length L is the number of bits, that it takes to communicate a model $\mathcal{M}$, the data D (in our case landmark positions) encoded with help of this model, and a residual error:

$$L(D,\mathcal{M}) = L(\mathcal{M}) + L(D|\mathcal{M}) + \mathcal{R} \tag{4.7}$$

The data term is associated with the reconstruction error, while the model term penalizes over-fitting through the use of expensive (in terms of number of parameters) models. In our case it provides information about the compactness of models describing the joint variation of the landmarks i and j, or equivalently about the redundancy in their position information in the training set. We expect low values for landmarks, that behave in a coherent way. That is, if $d_k(i,j)$ is the minimal description length [Davies *et al.*, 2002b], then the normalized graph Laplacian construction [Chung, 1997] allows us to construct a reversible Markov chain from the symmetric graph defined by the nodes and edges. We first consider

$$\begin{aligned} d(i) &= \sum_j k(i,j) \\ p(i,j) &= \frac{k(i,j)}{d(i)}, \end{aligned} \tag{4.8}$$

where ,

$$k(i,j) = e^{-\frac{d_k(i,j)}{\epsilon}}. \tag{4.9}$$

The fast decay ϵ represents a scale parameter, where $\epsilon > 0$. This parameter characterizes the notion of neighborhood between the landmarks, thus we can formulate x_i is ϵ-close to x_j. As shown in [Belkin and Niyogi, 2003], it is possible to opt for a kernel expressed as $\exp^{-\left(\frac{\|x_i - yj\|}{\epsilon}\right)^2}$ with a careful choice of ϵ, and then compute the normalized graph laplacian from this kernel, which corresponds to an approximation of the heat kernel.

The obtained kernel P conserves the positivity property, while loosing the symmetry, however now it satisfies:

$$\sum_j p(i,j) = 1 \tag{4.10}$$

Thus through Equation. (4.10), we can consider that the Markov chain is given by the non-symmetric matrix P with entries $p(i,j)$, and its powers P^t correspond to an increasing time in the chain, and to the according propagation of probabilities. Equivalently we can state that P represents the transition kernel of a Markov chain and it engenders a diffusion operator defined as

$$Pf(x) = \sum a(x,y) f(y) d\mu(y) \tag{4.11}$$

4.4.1.2 Diffusion Distance

The Markov chain captures the shape variation behavior by connecting groups of coherent landmarks with high-valued edges, while having low value edges between landmarks, that share only limited mutual information. The pairwise relation is captured by the according *diffusion distance*.

So if we recapitulate, we find that the kernel k describes the systematic behavior of the data while at the same time capturing some relevant geometric feature, while the Markov chain specifies the directions of the propagation following the kernel values. When we continue the random walk, the local geometry information keep propagating.

An eigenvalue analysis of P allows to generate a *diffusion map* [Coifman and Lafon, 2006], a metric space, in which a *diffusion distance* parameterized by t

$$D_t(i,j) = \sum_u \frac{(p_t(i,l) - p_t(j,l))^2}{\pi(l)} \tag{4.12}$$

$$\pi(i) = \frac{d(i)}{\sum_j d(j)} \tag{4.13}$$

As underlined by the authors in [Coifman and Lafon, 2006], the diffusion distance is smaller between points with many high value connections. The distance shows robustness towards noise, in contrast to the geodesic distance [Coifman and Maggioni, 2006].

4.4.1.3 Shape Maps

The eigenvalue decomposition of P leads to eigenvalues $\lambda_1, \lambda_2, \ldots$ and eigenfunctions $\Psi_1, \Psi_2, \ldots$ that satisfy

$$P\Psi_i = \lambda_i \Psi_i \tag{4.14}$$

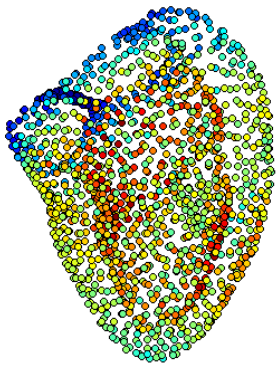


Figure 4.1 — Left Ventricle Shape Map estimation.

The diffusion map for the diffusion operator and t timesteps, $\Psi_t : X \longrightarrow \mathbb{R}^w$ embeds each node/landmark $i = 1, \ldots, m$, in the Markov chain into a w dimensional Euclidean space, resulting in the following family of diffusion maps $\{\Psi_t\}_{t \in \aleph}$

$$\Psi_t(i) \stackrel{\triangle}{=} \begin{pmatrix} \lambda_1^t \Psi_1(i) \\ \lambda_2^t \Psi_2(i) \\ \vdots \\ \lambda_w^t \Psi_w(i) \end{pmatrix} \tag{4.15}$$

where the diffusion distance defined in (4.12) becomes the Euclidean distance between the images of i and j, $\mathbf{\Psi}_t(i)$ and $\mathbf{\Psi}_t(j)$,

$$\|\mathbf{\Psi}_t(i) - \mathbf{\Psi}_t(j)\| = D_t(i,j). \tag{4.16}$$

Once the diffusion map is generated, the density estimation can be performed using an Euclidean approach. The i^{th} landmark has an image in the map, which we denote by $\Psi_i = \mathbf{\Psi}_t(i)$. The diffusion map $\mathcal{S}$, is a metric space and therefore we can estimate the density d_{Ψ_i} of the landmark images $\Psi_i \in \mathcal{S}$ for each point. The density relates to the number of landmarks, that can be encoded by the same model while retaining low description length. It is a measure of redundancy [Wallace and Dowe, 1999], since a model that represents the shape variation of a set of landmarks with images Ψ_i in a small neighborhood in $\mathcal{S}$ is compact - according to the generation of the Markov chain - and indicates that the mutual information that landmarks carry about each other is high.

We aim for a sampling, in which the information a landmark carry is more balanced. Each pair of landmarks should share an approximately equal amount of information about each other. An equal amount of mutual information about about other landmarks k, which have diffusion map images Ψ_k in its neighborhood. This would result in a uniform distribution of images Ψ_i in $\mathcal{S}$. In other words the shape map assigns each landmark a position Ψ_i. The distance between Ψ_i and Ψ_j in the shape map corresponds to the coherence of the behavior of landmarks in the training set. We aim at a uniform sampling in the shape map, so that the mutual information landmarks carry about each other is evenly distributed. The next step is to add appearance information to the map.

4.4.2 Image Support

The shape diffusion map represents the shape variation structure of the training examples. The appearance information, that is used during search, is not distributed evenly on the entire object, too. In the case of muscle MRIs only a small ratio of the surface carries distinctive appearance (see Fig. 4.2), that allows for a separation between background and foreground. To account for this variability we calculate the *image support* at each landmark position during training. If we can assign a value to a landmark relating to distinctive texture in the training set $\mathcal{V}$ we can further differentiate the representation $\hat{\mathcal{V}}$. Conceptually, the model should use landmarks with salient appearance for the inference

from the data, while reconstruction the remaining parts of the shape, as described in the previous section.

For each landmark i we denote by g_i the image support in the training set. g relates to the chosen search strategy. Since we employ local texture patches, we derive g based on the distinctiveness of the texture at the landmark position. In Fig. 4.3 the image support for calf muscles, and left ventricles is depicted. We calculate the correlation of texture appearance in the vicinity of landmark positions in the training set. For distinctive features, the correlation can be expected to show a peak at the correct position. Let $b(\mathbf{x}_i^j)$ be the learned texture patch at the correct landmark position $\mathbf{x}_i$ in the training example $\mathbf{I}_j$, and for landmark positions in a local neighborhood $\mathcal{N}$ let $Q_i^j(\mathbf{x})$ be the correlation between the patch $b(\mathbf{x})$ and $b(\mathbf{x}_i^j)$ normalized within the neighborhood, i.e.

$$\int_{\mathbf{x}\in\mathcal{N}} Q_i^j(\mathbf{x}) = 1, \tag{4.17}$$

then the image support is

$$g_i = mean_{j=1,\ldots,n}\left(\frac{Q_i^j(\mathbf{x})}{\int_{\mathbf{x}\in\mathcal{N}\setminus\mathbf{x}_i^j} Q_i^j(\mathbf{x})}\right). \tag{4.18}$$

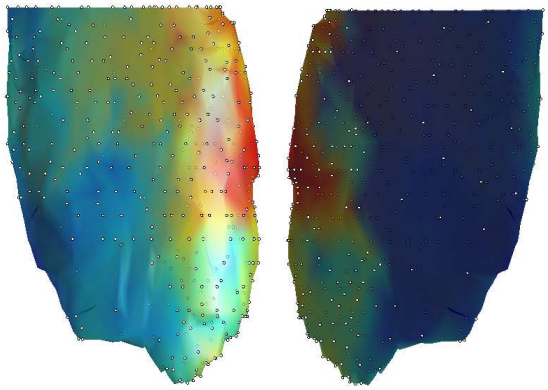
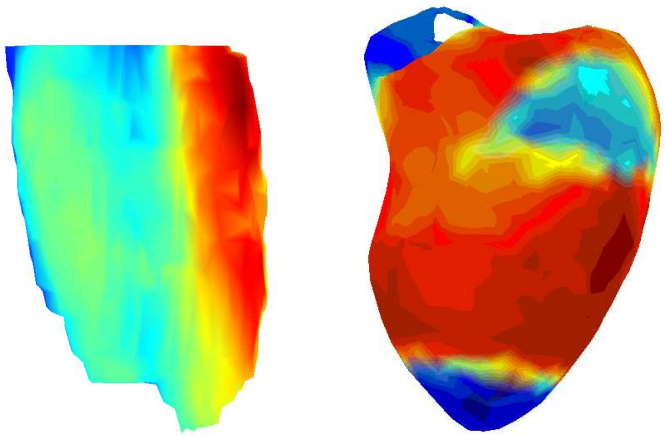


Figure 4.2 — Surface of a calf muscle: image support on the outer and inner part, and the sparse model points.

That is, for a landmark in $\mathcal{V}$ the image support is calculated from the local appearance behavior at the corresponding positions in the training set.



(a) (b)

Figure 4.3 — Image support: a. muscle surface, b. left ventricle.

In other words the choice of the landmarks will be established according to ;

- Equal amount of information with regard to the shape.
- Maximum amount of valuable image information.

The reconstruction of shapes from sparse set of K key landmarks chosen from a set of landmarks N, with $K << NL$.

4.5 Sparse Sampling of the Data

Given a metric space $\mathcal{S}$ that captures the statistical shape behavior, an according density d_{Ψ_i} and an image support g_i for each landmark we obtain a sparse sampling by minimizing the integral of absolute gradient value in the map $\mathcal{S}$

$$\mathcal{C}(\hat{\mathcal{V}}) = \int_{\mathcal{S}, i \in \hat{\mathcal{V}}} |\nabla(d_{\Psi_i}/\gamma g_i)|, \tag{4.19}$$

by choosing a subset of landmarks. That is, the function reaches a minimum if an even distribution of landmark images weighted by g_i is obtained in the shape map. This

distribution favors landmarks that have high image support in the training set, while integrating the statistical shape modeling and reconstruction properties of individual landmarks.

In practice, given a diffusion map $\mathcal{S}$, a set of object landmarks images $\mathcal{Y}_0 = \{\Psi_1, \dots \Psi_m\} \subset \mathcal{S}$, $\hat{\mathcal{Y}}_0 = \emptyset$, and the according densities $d_{\Psi_1}, \dots d_{\Psi_m}$, and a value r, we perform the sparse sampling in the following iterative way:

1. choose $i : d_{\Psi_i} = \max(\{d_{\Psi_j} : \Psi_j \in \mathcal{Y}_t\})$;
2. set $\mathcal{Y}_{t+1} = \mathcal{Y}_t \setminus \{\Psi_i \cup \Psi_j : \|\Psi_j - \Psi_i\| \leq r/\gamma g_i\}$, $\hat{\mathcal{Y}}_{t+1} = \hat{\mathcal{Y}}_t \cup \Psi_i$,

and iterate until $\mathcal{Y} = \emptyset$. The value r controls the mean density of the sparse representation.

This results in a set $\hat{\mathcal{Y}}$ and a corresponding set of landmarks $\hat{\mathcal{V}}$ that forms the sparse model representation, in which the mutual information between landmarks and the appearance information at landmark positions is distributed evenly.

In the following we will explain how to reconstruct the entire object $\mathbf{X}$ form the sparse representation $\hat{\mathbf{X}}$ utilizing the diffusion map $\mathcal{S}$. In Fig. 4.2 a sparse sampling for calf muscles is depicted together with the color coded image support.

4.6 Search Process

After the model is trained, we have a complete set of landmarks in the training set, and a set of optimally subsampled sparse landmarks. Each landmark corresponds to a local appearance in the training examples. In the following, we will first explain how to reconstruct the full set of landmarks from the sparse subset. With this, we will then outline the search procedure, that fits the model to new data, given the shape variation model, the reconstruction algorithm, and the appearance knowledge for each landmark.

4.6.1 Reconstruction and Missing values imputation

The reconstruction of the shape consists of inferring the positions of the entire shape $\mathbf{V_i} = \langle \mathbf{x}_1^i, \mathbf{x}_2^i, \dots \mathbf{x}_m^i \rangle$ from the sparse representation $\hat{\mathbf{V}}_\mathbf{i} = (\mathbf{x}_i)_{\Psi_i \in \hat{\mathcal{Y}}}$. We assume that

we model the shape variation locally by a multi variate Gaussian with axes along the principal components of the distribution. Furthermore, without loss of generality we can consider that we can derive a covariance matrix Σ for the position variation of each subset of landmarks in the shape after Procrustes alignment. The alignment discards the influence of global displacements of the local landmark configuration.

For a landmark $\mathbf{x}_i$ not in the sparse representation the reconstruction can be formulated in the following way: We choose the l nearest neighbors of Ψ_i in $\hat{\mathcal{Y}}$ (i.e. the landmarks, that exhibit the highest coherence of behavior in the training set and are part of the sparse representation). The use of the closest neighbors in the shape diffusion map ensures a reconstruction based on the functionally closest related landmarks as opposed to spatial neighbors.

The model learnt from the training data for this sub-set $\mathbf{V_{i,j}}'$ of landmarks or *reconstruction kernel* comprises a mean μ and a covariance matrix Σ. The shape vector is partitioned into the observed part $\mathbf{V_{i,j}}^a$ of the sparse representation and the missing part $\mathbf{V_{i,j}}^m$ (i.e. one or several missing landmarks). Accordingly we partition the covariance matrix into sub matrix corresponding to the observed values or coordinates Σ^{aa}, and the submatrix corresponding to the missing values Σ^{mm}, the submatrix describing their relation is $\Sigma^{am} = \Sigma^{ma\top}$, i.e.:

$$\Sigma = \begin{pmatrix} \Sigma^{aa} & \Sigma^{am} \\ \Sigma^{am\top} & \Sigma^{mm} \end{pmatrix}. \tag{4.20}$$

Now we can estimate the values of the remaining shape landmarks by a linear regression model:

$$\mathbf{V_{i,j}}^m = \mu^m + (\mathbf{V_{i,j}}^a - \mu^a)\mathbf{B} + e, \tag{4.21}$$

where

$$\mathbf{B} = \Sigma^{aa-1}\Sigma^{am}. \tag{4.22}$$

$\mathbf{B}$ is the regression matrix, $\mathbf{X}^m$ is the conditional maximum likelihood estimate of the missing part of the shape vector, and e is a residual error. See [Schneider, 2001] for a concise explanation of imputation. Therefore in the case of a linear model, the local sparse reconstruction function is given by

$$\mathbf{V'_{i,j}} \mapsto \begin{pmatrix} \mathbf{V'_{i,j}} \\ \\ \mathbf{V_{i,j}}^m \end{pmatrix} \tag{4.23}$$

It allows to reconstruct the entire object shape from the sparse representation, while using the relations learnt from the training set to define local reconstruction kernels in the map $\mathcal{S}$.

4.6.2 Inference from New Data

Let us now consider a new data set where the goal is to determine the position of the object being modeled. Such an inference process often involves the definition of an objective function that seeks for an admissible solution being supported from the observations. In a standard shape model inference approach, the positions of landmarks in new data are estimated by an energy minimization that involves both shape prior and appearance costs.

The search with the sparse model representation $\hat{\mathbf{V}}$, the according reconstruction function P, and the appearance models $(b_i)_i = 1, \ldots, N$ for each landmark is performed in an iterative manner. Based on a coarse initialization the landmark positions of $\hat{\mathbf{V}}$ are updated according to the appearance model. For each landmark the position with highest probability with regard to a local texture patch is chosen. Then the shape is constraint by either a local or global statistical shape constraint. In our work we use a multivariate Gaussian. However, alternatives, like spherical wavelets [Nain *et al.*, 2007], or elasticity based constraints [Taron *et al.*, 2007] can be utilized in a similar manner. After convergence the entire shape $\mathbf{V}$ is reconstructed from $\hat{\mathbf{V}}$ by the sparse reconstruction function P.

4.7 Experimental Validation

4.7.1 Experimental Set-up and Data Acquisition

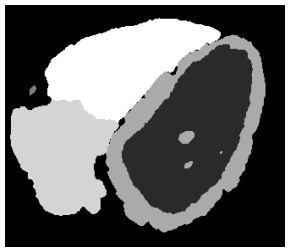
To evaluate the performance of the proposed method we report experiments on two data sets:

1. A set of 25 T1 weighted MRI calf muscles divided into two groups: 20 healthy control patients and 5 unhealthy cases. For each volume there are 90 slices of 4mm thickness, and with voxel spacing 0.7812x0.7812x4 mm acquired with a 1.5T Siemens scanner. Standard of reference annotation by experts for the Medial

Gastrocnemius (MG) muscle, was available (see Fig.(4.4.b)). Correspondences for 895 landmarks on the surfaces were obtained by an MDL based optimization [Langs *et al.*, 2007].

The data acquisitions were performed in the Henry Mondor University hospital on a 1.5T MRI, with a Siemens scanner using a body extremity coil positioned on both calfs. A *T1*-weighted spin echo sequence (TR: 500 ms; TE: 15 ms; slice thickness: 4 mm, acquisition time: 2 min 30 s) were performed in the axial plane to visualize the boundaries of the muscle and fatty infiltration. Thanks to our collaboration with the clinical experts we were able to consider a manual segmentation of healthy and non healthy muscle for each patient on anatomy images as presented through figure 4.4. This supervised ground truth segmentation is an essential step of pre processing to help the further shape analysis.

2. A set of 25 CT volumes of the heart, with an approximate voxel spacing of 1.5 mm, for which 90 anatomical standard of reference landmarks, and a set of 726 control points for the left ventricle was available, also with available ground truth segmentation from experts Fig.(4.4.a) concerning the diastole and the systole.



(a) Ground truth segmentation of papillary muscles

(b) *T-1* MRI slice supervised segmentation of a human calf

Figure 4.4 — Standard reference segmentation of respectively the calf muscle and the left ventricle

Normal volunteers for calf muscle study were from the MAS laboratory staff, whereas unhealthy cases were from Henry Mondor Universitary Hospital patients and more particularly from the radiology department. Please refer to Fig. 4.5 for an overview of the shape learning phase.

For both data sets we evaluated the reconstruction and search behavior of sparse models. To assess the shape representation of the sparse sampling, we sub-sampled the

shapes with landmarks either evenly distributed in the real space, or evenly distributed in the shape diffusion map, while neglecting appearance. The goal is to understand how the sparse sampling based on the density in the shape map affects the reconstruction of missing landmarks.

To evaluate the search behavior we compared sparse shape models with a standard shape model search in an active shape model manner, based on an even sampling of the object surface, and gradients in the volumes.

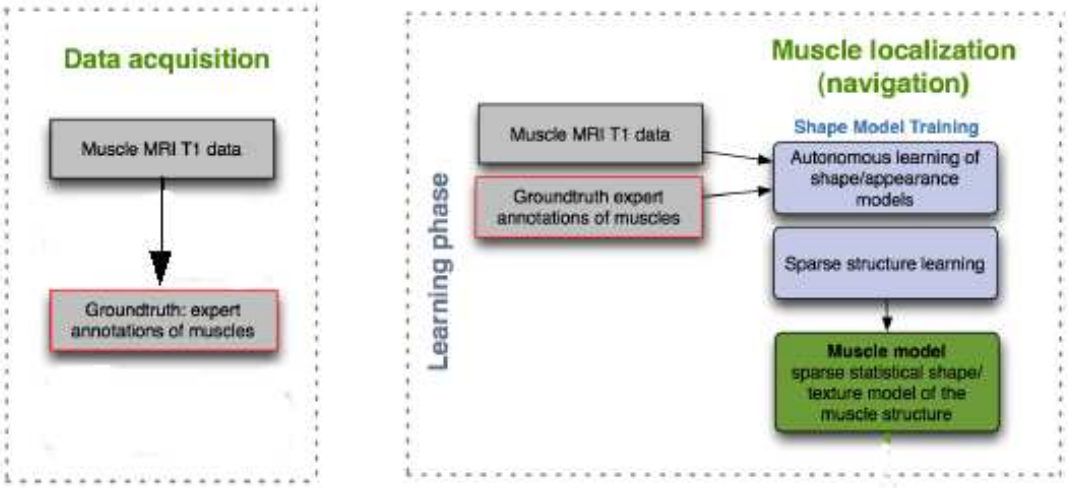


Figure 4.5 — Overview of the learning phase

4.7.2 Results

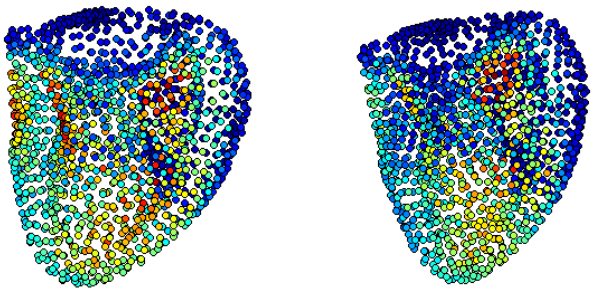
For both data sets (muscles, hearts) a sparse representation was built based on both shape model and image support. Models were initialized with minimal overlap to the target shape, and the accuracy of the final result was quantified by means of the mean landmark error between standard of reference annotation and search result. Indeed, the sparse model was able to recover the shape with superior accuracy.

In Tab. 4.1 mean landmark errors after search convergence for standard shape models, and sparse shape models are reported. In the muscle data the standard search approach failed due to the ambiguous texture in large regions of the target shape. In Fig. 4.8 examples for standard and sparse model search results are depicted. An interesting observation was that for calf muscle image support and diffusion map density gave com-

Mean Error	Heart Data	Calf Data
Standard Shape Model	18.92	37.78
Sparse Shape Model	7.85	9.51

***Table 4.1* —** Landmark error in voxel after finishing search with standard model, and sparse model respectively for heart and calf muscle data.

plementary distributions. That indicates, that it is worthwhile to use both informations for the representation building but raises the question of an appropriate weighting, and its dependence on the overall data variability. This will be subject of ongoing research on more exhaustive data sets. For the heart data, the search was initialized with minimum overlap. Standard search results in a mean error of 18.92 voxels, while sparse models obtain a mean landmark error of 7.58 voxel. An example of the resulting search performance with improved accuracy of sparse shape models on muscle data is shown in Fig. 4.11.



(a) (b)

***Figure 4.6* —** Shape Map estimation with different number of landmarks. The saturation encodes the density in the shape map indeed a high saturation stands for a high density.

In Fig. 4.7 shape diffusion maps, and densities are depicted for a set of calf muscles, and a set of left ventricles.

Reconstruction Reconstruction performance of sparse vs. standard models, with different noise ratios.

Search Search performance: convergence speed, and accuracy sparse vs. standard.

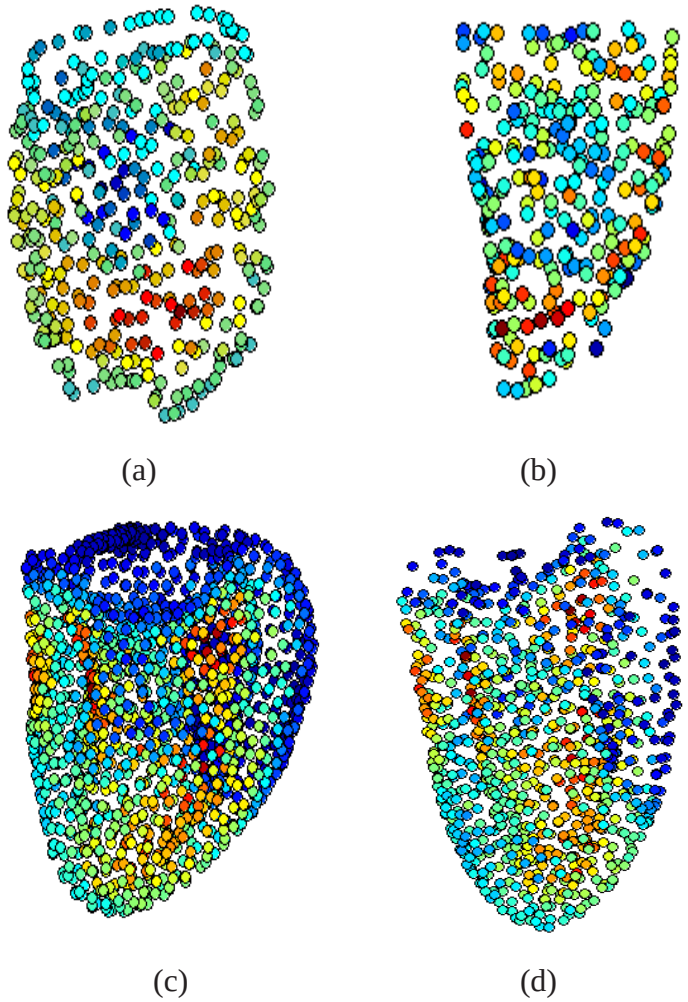


Figure 4.7 — Sparse shape models: Calf muscle: (a) color coded density in the shape diffusion map according sparse landmark distribution (b). Left ventricle: (c) shape diffusion map density and (d) sparse sub-sampling.

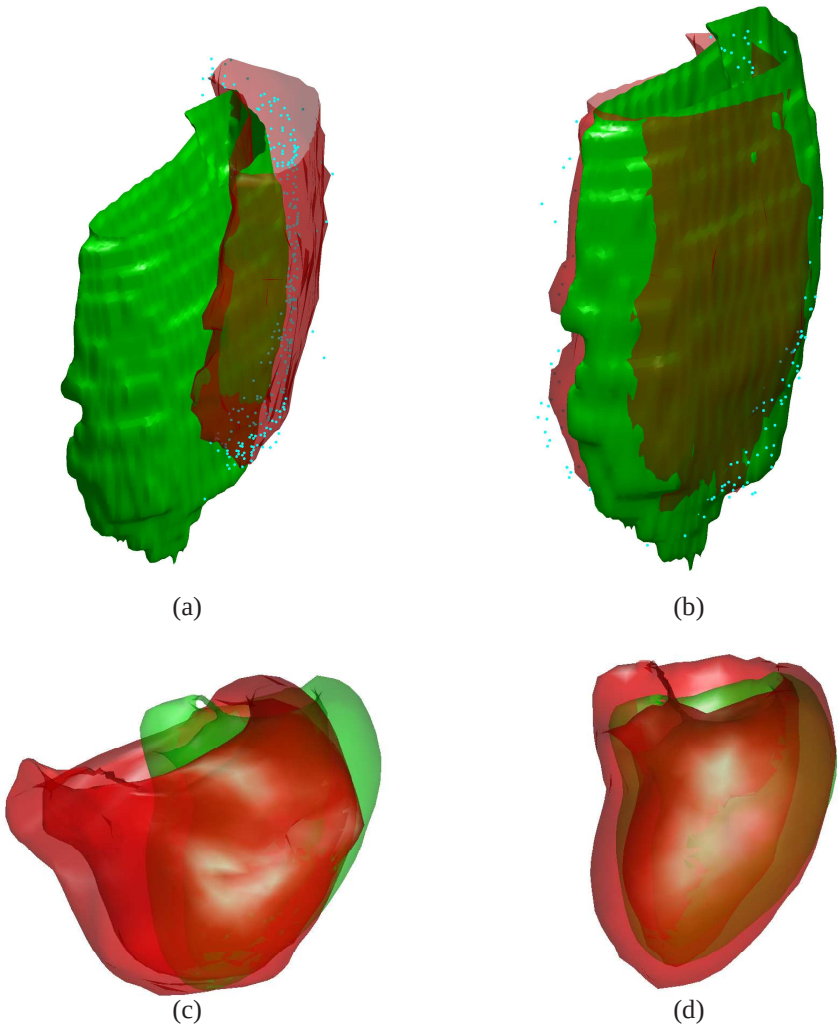
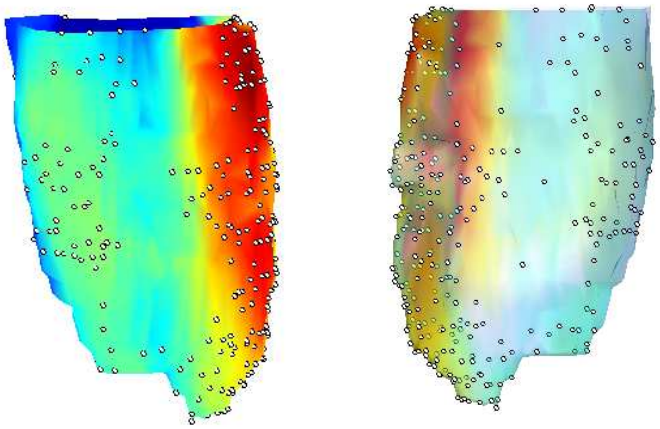


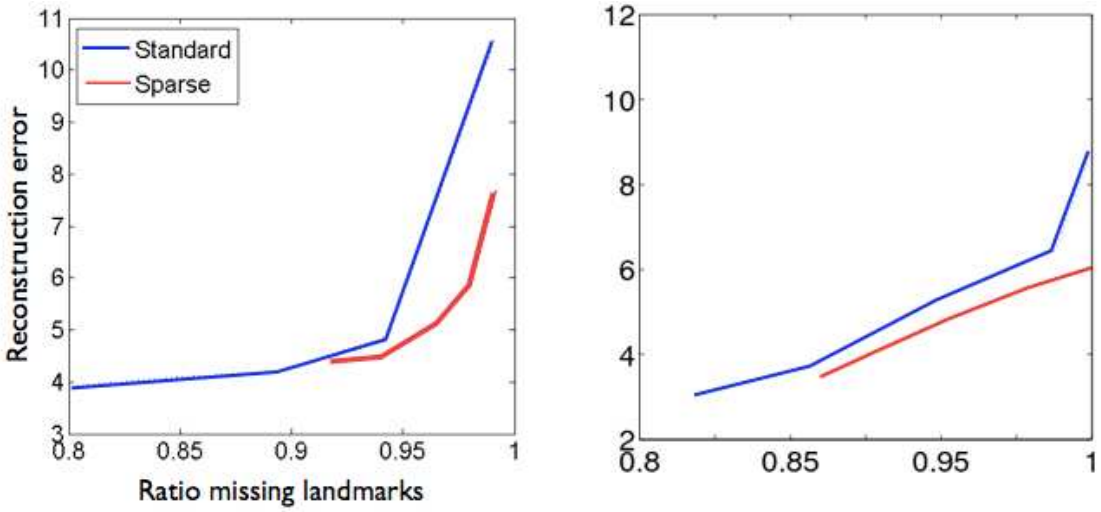
***Figure 4.8* —** Model search result for MRI calf (upper row) and heart muscle (lower row) data, green: standard of reference segmentation, red: search results for a. and c. standard gradient search approach, and uniform sampling, b. and d. sparse shape models.

Reconstruction results are shown in Fig. 4.10. With an equal ratio of missing landmarks (X-axis) the sparse sampling based on the shape diffusion map consistently outperforms uniform sub-sampling in the object space. The advantage becomes more pronounced with very high ratios. This indicates that a high amount of relevant information can be captured in a small sub-set of landmarks, when the modeling relations between them are considered by means of the shape diffusion map. As far as the heart is con-



(a) (b)

Figure 4.9 — Result of the sparse subsampling of the landmarks on the two sides of the MG calf muscle surface.



(a) (b)

Figure 4.10 — Reconstruction: accuracy of the shape reconstruction with different sparsity levels: uniform sub-sampling (in blue) vs. weighted sub-sampling (in red) on muscle and heart data.

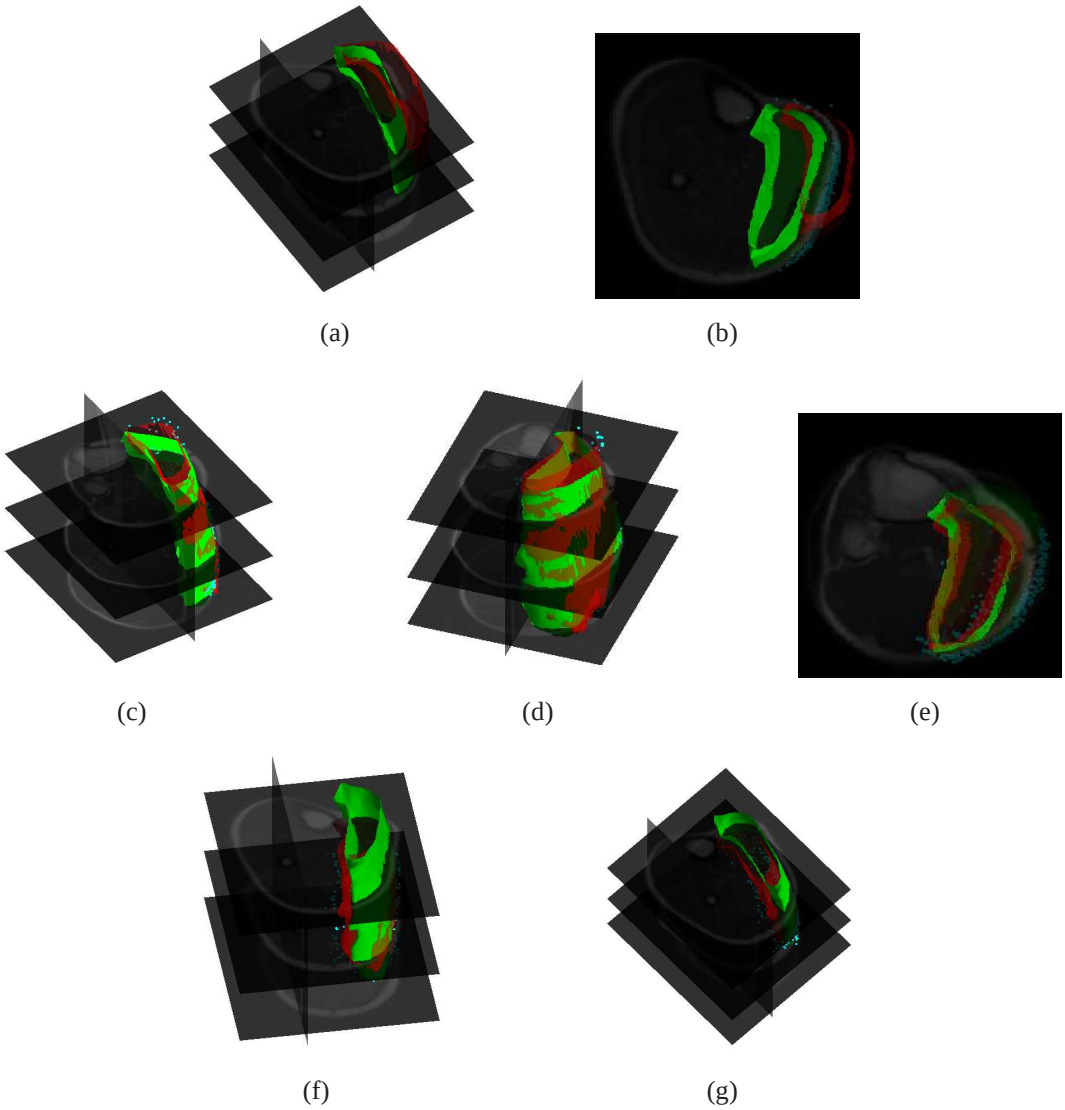


Figure 4.11 — Search: Model search result for MRI calf , green: standard of reference segmentation, red: search results. For (a) and (b) standard gradient search approach, and uniform sampling, for (c), (d), (e) and (f) sparse shape models. The two lines of sparse results correspond to two different control subjects.

	SPCA	RPM	SSM
Calf Muscle	30.24	21.78	9.51
Left Ventricle	17.07	12.33	7.85

Table 4.2 — Comparative table results between Sparse Shape Model, Sparse PCA and RPM a robust version of ASM over the landmark error between the ground truth and the reconstructed object relative to the two different data sets of calf and left ventricle muscle.

cerned, the sparse mean error reconstruction is equivalent to 8.7648, whereas the mean standard error is 18.9274.

Another practical suggestion of comparison to some robust versions of the ASM, would be the Robust Point Matching method (*RPM*) [Chui and Rangarajan, 2000], for which the code is available on line (http://noodle.med.yale.edu/ chui/tps-rpm.html) Such a comparison is beneficial to the method as it would demonstrate its real performance.

The table results Tab.4.2 clearly demonstrate that the sparse shape model provides a far better reconstruction quality than Sparse PCA. A larger study with different pathologies and the use of kernel PCA could improve the reconstruction of the shapes.

4.8 Contributions

We present a method to extract a sparse model of an anatomical object, from a set of training image data, each of which contains a set of landmark points. The novelty of the method consists in the way we sample the shape based on shape variability and image information distribution in the training set.

Initial experiments with atlas based segmentation, or image gradient guided muscle separation showed insufficient accuracy, and fail in a majority of the cases. For this reason we have developed sparse shape models with the following properties: Sparse models use a priori knowledge about the shape and appearance of muscles, learned during a training phase. They learn the sparse distribution of reliable image content (high contrast between fat and muscle tissue) and integrate this knowledge in the representation of the muscle model; they also learn the intrinsic structure of the muscle shape variation and use it to parameterize the muscle shapes in an optimal way; finally they identify positions on the muscle surface in a repeatable way. That is, after learning the

model from a training set of muscles, corresponding positions can be localized in MRI data.

Our approach seeks to identify a set of points that have distinctive appearance while at the same time covering the object evenly with regard to the information contained in each sampling point, i.e. mutual information with its neighboring points. Toward fulfilling this goal, we have to first construct a diffusion map of the landmarks, a metric space that determines the mutual information between the landmarks. Starting from this map, the density at each landmark is computed in an Euclidean metric. A subset of the original landmarks is chosen so that the densities at the sampled landmarks are even by minimizing the integral of density gradient of the landmark set. To favor landmarks with salient appearance, each landmark is additionally weighted by the image support measure, the correlation of texture appearance in the vicinity of the landmark position, before the selection is done.

To reconstruct the full shape from the sparse sampling, each missing landmark is reconstructed in a linear regression manner from these neighbor points. We demonstrate the utilization of the sparse model in segmenting anatomical data sets - the calf muscle and left ventricle- and compare results with standard uniform sub-sampling of the object. With the same rate of sub sampling, the proposed model outperforms uniform sub-sampling of the object space both in terms of reconstruction error, and the performance during object search in new data.

The Sparse Shape Models adapt the representation to the shape variation and its redundancy, and to the local image support as observed in the training data. In that way it provides for an optimal landmarks set, and a mechanism to reconstruct the entire shape from this sub-set. Sparse Models also offer improved representative power of landmarks, and a better search stability and accuracy.

The proposed approach takes both the behavior of the landmarks in the training set, and the local appearance of the anatomical structure into account to obtain an optimal model. Our approach lies in taking into account both the geometric relationship and appearance during the sampling process. The use of appearance information, more precisely the idea of using landmarks with salient appearance (or texture distinctiveness), estimated in a training sample via a correlation function between local patches in a neighborhood and the sparse sampling of the data using our criteria as expressed in Equation. (4.19) is also of interest for the segmentation of other data that exhibits a heterogeneous distribution of informative image content. The experiments show the

advantages proposed over uniform subsampling method.

In the following chapter we are going to extend our approach a step further, and instead of only using the concept of diffusion to sub-sample the data, we will introduce a parameterization of the shapes based on the topology encoded in the diffusion maps learned during training.

CHAPTER

5 Hierarchical 3D Diffusion Wavelet Shape Priors

"Il faut toujours viser la lune, car même en cas d'échec on attérrit dans les étoiles."

Oscar Wild

5.1 Introduction

The optimal representation of the significant variations in a set of shapes at various scales is a challenging problem. Parameterizations chosen a priori pose limitations on the representative power, and introduce a bias on the model. An example is a reference manifold like a sphere that is used to parameterize shapes of genus zero.

In the following chapter we introduce a more general formulation of a decomposable shape parameterization. It can adapt to the training population of shapes. It allows to learn an optimal parameterization from the training set, and to represent the shape variation in this reference frame. Furthermore the shape parameters deal with global and local variations present in the population separately. The representation is based on the wavelet transform and uses its ability to exploit the intrinsic multi-scale nature of the data.

We explore a method for the parameterization of the shape variation observed in the training data by means of diffusion wavelets [Coifman and Lafon, 2006]. the diffusion wavelet is a very generic construction based on the notion of a diffusion kernel. It can be viewed as a generalization of standard parameterizations, e.g.: the kernel for a triangulated spherical surface would be the adjacency matrix weighted by the mutual distances. Defining the topology by a diffusion kernel instead of a fixed genus-zero manifold allows us to incorporate, and even to learn, complex interaction patterns observed in the training data, and use them to build an efficient shape variation prior. In the following we will first outline the basics of diffusion wavelets [Coifman and Maggioni, 2006], then explain the associated shape variation representation, and finally detail how the orthomax principle [Kaiser, 1958] can be used to separate coherent sub-regions of the shape.

In this chapter we will introduce the use of diffusion wavelets to represent the variation of shapes. However, instead of relying on a pre-defined manifold (e.g. a sphere [Nain *et al.*, 2007]) we will learn the topology of the wavelet domain from the training data, and will encode it in a diffusion kernel. The kernel allows us to learn and define arbitrary wavelet hierarchies, and thus to make optimal use of the training data. The wavelet representation used in this thesis is based on *diffusion wavelets* proposed in [Coifman and Maggioni, 2006].

Let's recall that the diffusion distance concept was already introduced and used in the last chapter, and will also be exploited in the current one. The intuition behind such an approach is to seek a complete embedding of the landmarks in a Riemannian manifold.

It can also be considered as a wavelet equivalent of computing the eigenfunctions of the Laplace-Beltrami operator on an arbitrary manifold.

The remainder of the chapter is organized as follows: We will first present the basic of wavelets in Section. 5.2, and the related work in which they are applied to medical data in Section.5.3. We then detail the theoretical background concerning diffusion wavelets as well as variation modeling based on the shape representation using diffusion wavelets in Section. 5.4. In Section. 5.5 we focus on the manifold construction and the inference in new data. Finally in Section. 5.6 we report experimental results and a quantitative validation.

5.2 Wavelets

Wavelets are a robust mathematical tool for the hierarchical decomposition of functions. The theory is described extensively in e.g. [Meyer, 1993] [Mallat, 1989]. The decomposition allows for a representation in terms of a coarse overall shape, that is enriched by details in a coarse to fine hierarchy. Starting from the definition of a mother wavelet as a localized function on the plane, on which it is possible to carry on affine transformation, i.e translation, rotation and dilation. The wavelet coefficients construction is the result of the correlation of the signal with the transformed version of the mother wavelet, in other words this is to be considered as wavelet transform of a signal on the plane. Unlike principal components analysis and Fourier basis functions that represent global shape descriptors, wavelets captures both global and local variations.

The wavelets provide so an elegant technique for representing detail levels regardless of the interest function type (e.g. images, curves, surfaces). Thanks to their local support in both space and frequency, wavelets are suited for sparse functions approximations. Their major strengths are the compact support of basis functions as well as the inherently hierarchical representation. The domain upon which the wavelet hierarchy is defined is of prime importance for their representative power. Roughly speaking the theorem introduced by [Meyer, 1993] [Mallat, 1989] establishes that given an orthogonal multiresolution analysis, one can find a function whose dilates and translates will generate an orthonormal basis. In practice we can find several wavelet transform, such as the Discrete Wavelet Transform (DWT) or rather the Fast Wavelet Transform (FWT).

On the other hand, wavelets can be considered as a rotation transform in function

space. As far as compression is concerned wavelets are designed to accomplish rotations leading to decorrelate image data by using vanishing moments. Wavelet coefficients close to zero can consequently be removed without loosing much information through the reconstruction.

It is beyond the scope of this section to give the detailed mathematical definitions about the different wavelet transform discussed in the literature. However we will above all focus in this chapter on spherical, harmonic and diffusion wavelets.

5.3 Previous Work

The wavelet representation offers a wide range of advantages for the processing of different medical imaging modalities, such as PET, FMRI, MRI, CT... These first generation of wavelets have contributed to various applications, among the most famous image compression, denoising and retrieval. However during the two last decades a substantial amount of work has been published concerning shape modeling based on wavelet representation.

[Brammer, 1998] presented a multidimensional wavelet analysis of functional magnetic resonance images, whereas [Turkheimer *et al.*, 2006] utilized wavelets for multiresolution Bayesian regression in PET dynamic studies. In [Nowak, 1999] wavelet-based Rician noise is removed for magnetic resonance imaging. In 1999 [Wolstenholme and Taylor, 1999] introduced the idea of modelling wavelet coefficient in an Active Appearance Models framework by the means of Haar wavelet, an issue extended afterward by [Stegmann *et al.*, 2004]. During the latter work wavelets were incorporated to AAM to obtain thus a major decrease in storage requirements. The authors were already anticipating that a wavelet based method will play an important role in medical imaging analysis.

The first work closely related with our approach was presented by [Davatzikos *et al.*, 2003]. The idea there was to build a hierarchical active shape models of 2-D anatomical objects using 1-D wavelets, which are then used for shape based image segmentation. The wavelet transform is interesting in a way that it is considered as a decorrelator of real-world signals [Mallat, 1989] and thus the covariance matrix of the wavelet coefficients is sparse. Starting from this assumption, the covariance matrix of the wavelet coefficients is estimated as a a block diagonal matrix, when reordering the coefficients

in the right range. A diagonal block of the covariance matrix would be formed by the coefficients appertaining to the same band. The authors use a logarithm tree for the coefficients bands grouping in order to divide the space-frequency domain. In consequence the coefficients holding the same scale and neighboring spatial location would be positioned in the same band, following the inherited assumption that once coefficients are neighbors in space and scale implies that they are exhaustively correlated. The wavelet transform is therefore exploited to reorganize the model into a hierarchy of several parts: the lower bands of the transform represent the global shape variations, whereas the higher bands correspond to more local changes. Every band is modeled independently from the rest. The philosophy behind the use of wavelet decomposition is to improve the shape model flexibility by dividing the latter into independent components.

A further extension for 3D shapes was introduced by [Yu *et al.*, 2007] and [Nain *et al.*, 2007]. The generalization of wavelets to the sphere is not straightforward, particularly due to the issues of sampling and dilation [Vandergheynst and Wiaux, 2010]. The use of spherical wavelets introduced in [Schröder and Sweldens, 1995] is therefore considered as second wavelet generation. This type of wavelet came along to process closed genus-zero surfaces to reveal their shape characteristics. Especially if one has to project spherical data into the Euclidean space it may sometimes lead to severe distortion.

[Nain *et al.*, 2007] proposes an improvement of the work of [Davatzikos *et al.*, 2003] by implementing a multi-scale wavelet based segmentation for 3D medical shapes using conformal mapping and a subsequent spherical wavelet representation. In addition they extend the work by proposing a novel algorithm to discover optimal multi-scale bands from the data.

More practically the authors use spherical wavelets, a decomposition of meshes topologically equivalent to spheres, to analyze the training set and represent it with a wavelet basis. Afterwards the coefficients with very small variance are discarded, and the others gathered into bands. Later on principal component analysis estimates the probability density function relative for each band, and an optimization is carried on to match the model to new volume. Subsequently we end up with a more reduced model that prevails the active shape model. Spherical wavelet can so characterize shape variation in a local fashion in both space and frequency, in contrast to spherical harmonics that have a global basis set.

As a matter of fact the authors also claim that comparing to [Davatzikos *et al.*, 2003] who cluster coefficients of spatially adjacent bases into bands in each frequency plane,

they proceed to the clustering of highly correlated coefficients into a band, while at the same time applying a constraint across bands having minimum cross-correlation [Nain, 2006].

A parallel study was conducted by [Yu *et al.*, 2007] who compared the efficiency of Spherical Harmonics (*SPHARM*) and spherical wavelets in extracting local shape variations of neuroimaging data. Spherical wavelet proved to be more effective concerning computational time.

However already in [Schröder and Sweldens, 1995] the authors speculate on a generalization of their work on wavelets and spheres towards arbitrary topologies.

Specificities and Limitations of Spherical Harmonic and Wavelet Being a key issue, the choice of the wavelet type is determinant, which can principally be related to the nature of data, however we will mainly concentrate on some shortcoming related to harmonic and spherical wavelets:

Spherical Harmonic Descriptors SPHARM [Yu *et al.*, 2007]: They can be considered as a natural extension of Fourier Transform on a sphere, and thus proved to be adequate for describing closed surface with spherical topology, which is not the case for our data. The coefficients in the spherical harmonic basis of different levels have the advantage to offer a mea- sure of the spatial frequency constituents that comprise the structure. Still they suffer from the limited global support.

Spherical Wavelet Unlike SPHARM approach that only depict global variations, Spherical wavelet functions can address this disadvantage by their local support at multiple resolutions levels, and this throughout using even fewer coefficients [Yu *et al.*, 2007]. Moreover one of the pre-processing required by spherical wavelet segmentation is the triangulation of the data and conformal mapping of the surfaces with spherical topology. Seeing that we are handling arbitrary manifolds (e.g multiple muscles) for which we aim to represent the shape variation, an automatic triangulation is not straightforward. Indeed instead of establishing a triangulation we will rather construct a kernel diffusion/operator.

These wavelet seems to be more adapted to data that can be expressed in the spherical coordinates, whereas in the ideal case one would like to use wavelets on surfaces of genus higher than zero. To overcome these limitations we will rather focus on a generalization

of classic wavelet based on a diffusion scheme, on which we will concentrate in the following section.

After the construction of wavelet around a sphere the authors in [Schröder and Sweldens, 1995] assume in their future work that the generalization to arbitrary topologies would be of great interest and success, which finally came through diffusion wavelets a decade later.

Most recently, and in parallel to our work [Essafi *et al.*, 2009b], a promising approach to use diffusion wavelet was published by [Zhu *et al.*, 2009]. The authors consider the idea to exploit diffusion wavelets for the matching of 3D shapes. There are substantial differences both in terms of method and scope to our work. In contrary to [Zhu *et al.*, 2009], we are aiming to build a generative model of shape variation, which adapts to the topology of a set training examples, and use it for the segmentation and reconstruction in new volume data. Instead they only employ diffusion wavelets as descriptor of shapes. On top of that, an important advantage of our method is the use of orthomax to obtain an optimal subdivision of the shape. Hierarchical shape models based on diffusion wavelets that adapt to arbitrary topologies have not yet been published, and their use in computer vision and medical image analysis has not been described.

A comprehensive overview of the wavelets transform on manifold can be found in [Antoine and Vandergheynst, 2009] and in the exhaustive book chapter [Schclar, 2008].

5.4 Diffusion Wavelet Model Construction

5.4.1 A Diffusion Operator Reflecting The Topology

We represent the shapes by a finite set of landmarks. For m landmarks the positions, $\mathbf{V}_i = \{\mathbf{x}_1^i, \mathbf{x}_2^i, \ldots \mathbf{x}_m^i, \}$, are known in N training images $\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_N$. That is, our shape knowledge comprises $\mathcal{V} = \{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_N\}$, where $\mathbf{x}_j^i \in \mathbb{R}^d$, and we call $\mathbf{V}_i \in \mathbb{R}^{d \times m}$ a shape.

Since we are only interested in the non-rigid deformation, all anatomical shapes are aligned by Procrustes analysis [Luo and Hancock, 2002], which produces the series of examples V_i^p, from which we compute the mean shape $\bar{V}^p$. After the registration, we can represent the shapes by their deviation S_i (Equation. (5.1)) from the mean shape,

$$S_i = V_i^p - \bar{V}^p \tag{5.1}$$

where

$$\bar{V}^p = \frac{1}{N}\sum_{i=1}^{N} V_i^p \tag{5.2}$$

Now we define a topology on the set of landmarks. The representation is based on a framework for multi-scale geometric graph analysis proposed in [Coifman *et al.*, 2005]. It applies the concept of diffusion to capture mutual relations between nodes in a Markov chain, that encodes the global neighborhood or modeling structure of the shape landmarks. In our case, this structure is the neighborhood relation between landmarks of the shape. It determines the domain upon which the wavelet representation is built. Diffusion maps provide a canonical representation of high-dimensional data. They allow us to encode spatial relations, or the behavior [Langs and Paragios, 2008] of the shape training population. The structure is encoded in a diffusion operator $T \in \mathbb{R}^{m\times m}$. This operator is then used to define the diffusion wavelets which represent both the global and local properties of the data in relation to the operator.

We build the diffusion operator T on the set of points embedded in a metric space in two different ways:

(a) Their mutual distance in the mean shape, or

(b) Their joint modeling behavior.

In the first case, and in order to build a matrix of graph weights for the points, we construct a local Gaussian kernel function centered at each point and then normalize the weight matrix through the symmetric Laplace-Beltrami to form the diffusion operator T. In the second case, when modeling their joint behavior, we derive the diffusion operator by probing the behavior of small subsets of the landmark set, according to the method described in [Langs and Paragios, 2008]. The resulting operator T reflects all pairwise relations or *neighborhoods* between individual points in the shape set.

As described in Chapter. 4, we define a diffusion operator T based on these measures. T is a self-adjoint operator conjugate to the Markov matrix $\mathbf{P}$ and $\mathbf{I} - T$ is the Laplacian on our data [Coifman and Maggioni, 2006].

Let's consider the graph G as described in the Section. 3.5. Let's also recall that we have $d_i = \sum_j k(i,j)$ as the total connection of the landmark i with the rest of data set, $\mathbf{D}$ corresponds to a diagonal matrix with d on its diagonal, and $\mathbf{P} = p_{ij}$ which origin is explained in Section. 4.4.1.1

$$Pf(x) = \sum a(x,y)f(y)d\mu(y) \tag{5.3}$$

Consequently and as shown in [Shi and Malik, 2000] the operator L constitutes the normalized Laplacian on the graph

$$L = \mathbf{D}^{-1/2}(\mathbf{I} - \mathbf{P})\mathbf{D}^{-1/2} \tag{5.4}$$

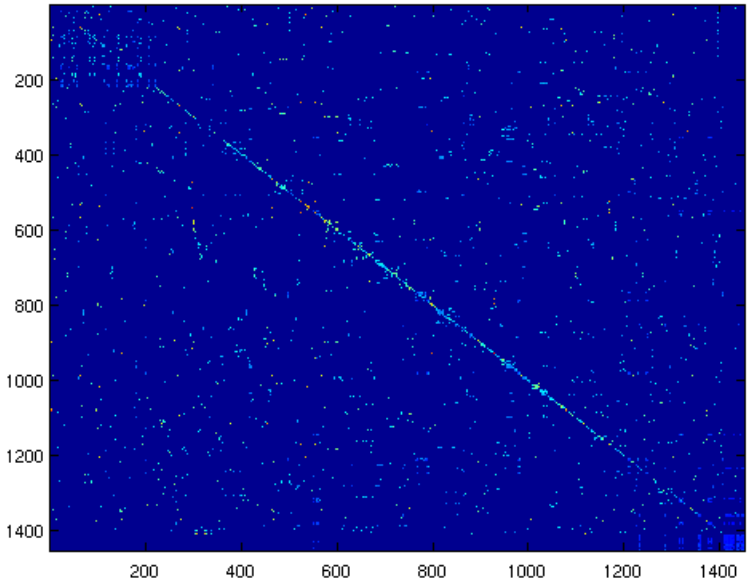


Figure 5.1 — The diffusion kernel operator T

For the remaining work we use a Laplace-Beltrami operator. We should remind here that the Laplacian operator captures the local geometry in the Euclidean geometry. The Laplace-Beltrami operator does the same, but also on Riemannian geometries i.e. manifolds. This corresponds to our case, because the landmarks are situated on a manifold defined either by their mutual euclidean distance or by the shape map property model.

More intuitively the Laplace-Beltrami operator can be seen as a generalization of the Euclidean representation of the Laplace operator to an arbitrary Riemannian manifold.

5.4.2 Diffusion Wavelets

In [Coifman and Maggioni, 2006] the authors generalize the idea of basic wavelets presented in [Meyer, 1993] [Mallat, 1989] by adjusting the wavelet construction to the geometry of the operator. One of the first assumptions of the authors is that the diffusion operator T is self adjoint and that it could be considered on the basis Φ_0 as following:

$$\Phi_0 = \{\delta_k\}_{k \in X} \tag{5.5}$$

where δ_k is the Dirac δ-function, and then take into consideration $\tilde{\Phi}_1 = \{T\delta_k\}_{k \in X}$ and build the wavelet in a multi resolution way as illustrated in Figure. 5.2.

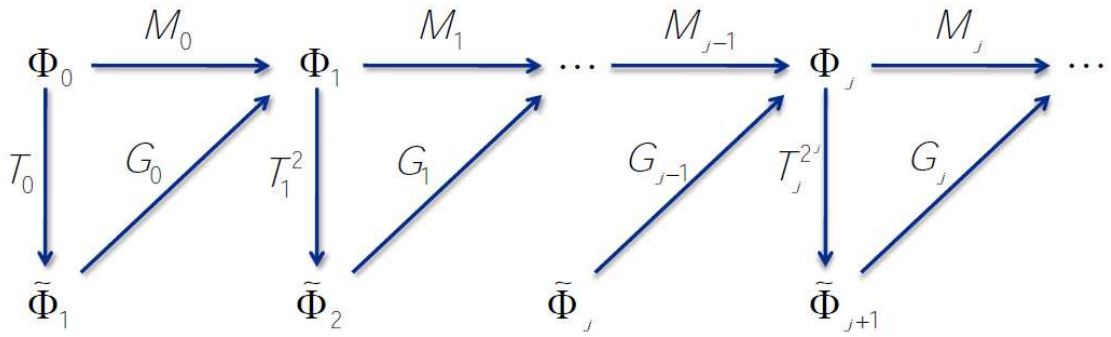


Figure 5.2 — Diffusion wavelet generation, downsampling and orthogonalization [Coifman and Maggioni, 2006] .

If we refer to the diagram in Figure. 5.2, one has to note that all triangles are commutative by construction and also that for any scale j there are some relations that have to be fulfilled.

$$M_j = G_j o T^{2j} \tag{5.6}$$

$$\tilde{\Phi}_j = T^{2j}\Phi_{j-1} \tag{5.7}$$

We have so a linear mapping from $\tilde{\Phi}_j$ to Φ_j. Let's note that G_j represent the local multi-scale Gram-Schmidt procedure [Daniel *et al.*, 1976], which ensures that G is not only linear and local but also sparse, consequently the orthonormal basis Φ_{j+1} would be coarser than Φ_j.

To allow a fast computation for the diffusion [Coifman and Maggioni, 2006] inspired by some physics and geometry examples, propose to work with the dyadic powers T^{2j} as a compressed representation of the operator. In fact those powers are known to decrease in rank, consequently they can lead to the function compression.

In Figure. 5.2 we can notice that first one has to apply T to a space of test functions at the finest scale, compress the range via a local orthonormalization procedure, represent T in the compressed range and compute T^2 on this range, compress its range and orthonormalize, and repeat this loop in a way that in the level j we are computing T^{2j}.

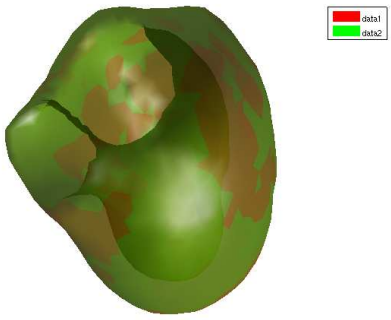


Figure 5.3 — Heart left ventricle: Comparison of a Ground truth and the Reconstruction by means of diffusion wavelets.

5.4.3 Shape Variation Modeling with Diffusion Wavelets

Given the diffusion operator T defining the manifold, we use the corresponding hierarchical diffusion wavelets, to represent the shape variation. First we build a hierarchical wavelet structure, the *diffusion wavelet tree*: We call upon a general multi resolution construction for efficiently computing, representing and compressing T^{2j}, for $j > 0$. The latter are dyadic powers of T, and we use them as dilation operators to move from one level to the next. We can expect it to be easier to compress high orders of the diffusion operator as they are supposed to be low ranked. During the down-sampling process,

and throughout a recursive sparse QR decomposition we obtain the orthonormal bases of scaling functions, $\Phi = \{\phi_j\}$ that represent a smooth bump function at the scale j, the wavelets Ψ_j, and compressed representation of T^{2j} on ϕ_j, for j in the requested range. The scaling functions $\{\phi_j\}$ and the orthogonal wavelets $\{\Psi_j\}$ are spanning the spaces V_j and W_j, where the first encompasses the coarse details. Both presented spaces exhibit the following proprieties: in (5.8) where the subspace W_j represent the orthogonal complement of the subspace V_{j+1} in V_j

$$\begin{aligned} V_{j+1} &\subseteq V_j \\ V_{j+1} &= V_j \oplus^{\perp} W_j \end{aligned} \tag{5.8}$$

Giving K as maximum number of levels to compute, we obtain a representation of T^{2j} onto a basis ϕ_j, with $1 \leq j \leq K$ after K steps. For a detailed description of this construction we refer the reader to [Coifman and Maggioni, 2006].

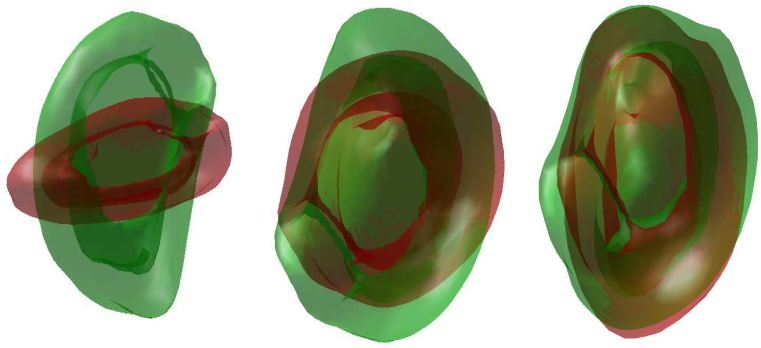
After building the diffusion wavelet tree Φ, we use it to represent the individual training shapes. We calculate the diffusion wavelet coefficient Γ on the deviation S_i from the mean of the aligned shapes, and obtain the following diffusion wavelet coefficients for an example S_i,

$$\Gamma_{S_i} = \Phi^{-1} S_i \tag{5.9}$$

Thus, the shape can be reconstructed by:

$$V_i^p = \bar{V}^p + \Phi \Gamma_{S_i} \tag{5.10}$$

Once we have generated the diffusion wavelet coefficients for all training examples, we build a model of the variation by means of the orthomax criterion - which will be described in the upcoming section - at each level. In the lowest level the coefficients provide information for a coarse approximation, whereas localized variations are captured by the higher-level coefficients according to the hierarchy. For each level j, with ($1 \leq j \leq K$), we consider $\Gamma_{level\,j}$ (Equation. (5.11)) and perform principle component analysis to reduce the dimension of the coefficient representation for all coefficients scales.



(a) (b) (c)

Figure 5.4 — Reconstructed surfaces for Heart CT data using projected wavelet coefficients on the set of principal components that represent 99% of the total variance at level 1. The axial view surfaces represent the $\pm 3\sqrt{(}\lambda_i)$ from left to right.

$$\Gamma_{level\,j} = \left\{\Gamma_{S_i/level=j}\right\}_{i=1\ldots N}, \tag{5.11}$$

This results in the eigenvectors $\Sigma = \{\sigma_j\}_{j=1\ldots K}$, the corresponding eigenvalues $\Lambda = \{\lambda_j\}_{j=1\ldots K}$ of the covariance matrix of the diffusion wavelets coefficients at each level j, and the according coefficients $\Gamma^*_{level_j}$ that represent each training shape in this coordinate system.

Consequently in each level the coefficients are expressed such as:

$$\Gamma_{level\,j} = \bar{\Gamma}_{level_j} + \sigma_j \left(\sigma'_j.\Gamma^*_{level_j}\right) \tag{5.12}$$

Based on the model parameters $\{\Lambda, \Sigma\}$ we can reconstruct a shape by first obtaining the diffusion wavelet coefficients $\Gamma_{S_i\,Rec}$ in each level, and then reconstructing the shape based on the diffusion wavelet tree:

$$V_i^p = \bar{V}^p + \Phi\Gamma_{S_i\,Rec} \tag{5.13}$$

This shape representation can now be used to model the shape variation in the training set, and represent it in the diffusion wavelet coefficient domain.

5.5 Prior Manifold Construction & Image-based Inference

We can consider different dimensionality reduction techniques like Principal Component Analysis, Linear Discriminant Analysis, Non Negative Matrix Factoriztion, statistical approximation methods like mixture models, Expectation Maximization, or recent spectral kernel methods like Locally Linear Embedding [Roweis and Saul, 2000] and Laplacian Eigenmaps [Belkin and Niyogi, 2003] in order to decrease the dimensionality of the diffusion wavelet coefficient representation of a training set of shapes. However given the ability of the diffusion model to capture relevant non-linear variations, we choose a simple dimensionality reduction technique, and a linear sub-space representation for our experiments. To obtain a sparse and localizad representation, we adopt the orthomax criterion [Kaiser, 1958].

5.5.1 The Orthomax Rotation

The orthomax criterion [Harman, 1976] is a technique belonging to the family of factor analysis, that allows to obtain a simple and compact hierarchical representation through a rotation of the model parameter system. This parametrization rotates the PCA modes so as to enhance sparsity, while at the same time preserving the orthogonality of components.

Despite of the simplicity of the orthomax criterion in terms of computation or conception, the shape analysis community in general and medical imaging literature in particular are devoid of studies carried on with this approach as stressed by [Stegmann *et al.*, 2006].

Orthomax rotations reflect a re-parameterization of the PCA space resulting in a simple basis. In [Stegmann *et al.*, 2006] [Leung and Bosch, 2007a] the orthomax was exploited as a straightforward method to select sparser modes. Alternatively [Sjostrand *et al.*, 2007] introduced sparse PCA where the orthogonality is not elementary, but on the meanwhile the method still brings out near-orthogonal components, whereas suffring from computational issues. It can also be considered as a natural continuation of the work on Independent Component Analysis (ICA) [Üzümcü *et al.*, 2003] aforementioned in section 4.2. as ICA does not imply orthogonality criteria but still gather the maximum of sparsity. In the same context the orthogonal orthomax criterion can be accounted

Orthomax Type	ω
Quartimax	0
Varimax	1
Equamax	$k/2$
Parsimax	$p\frac{k-1}{p+k-2}$

Table 5.1 — Variants of Orthomax Rotation [Harman, 1976]

for equivalent to the Crawford-Ferguson criterion, which is a weighted sum of row and column complexity of the eigenvector matrix [Leung and Bosch, 2007a] .

Let $\boldsymbol{R}$ be an orthonormal rotation matrix in $\mathbb{R}^k \times k$ where $\boldsymbol{R}_{i,j}$ represents the elements, and where k implies the number of eigenvectors with the largest eigenvalues $\lambda_{ii=1,..,k}$. Σ denotes as previously in the chapter the $p \times N$ eigenvectors matrix.

The orthogonal orthomax rotation matrix $\boldsymbol{R}$ is then calculated as follow:

$$\boldsymbol{R} = \arg\max_{R} \left(\sum_{j=1}^{k} \sum_{i=1}^{p} (\Sigma \boldsymbol{R})_{ij}^{4} - \frac{\omega}{p} \sum_{j=1}^{k} \left(\sum_{i=1}^{p} (\Sigma \boldsymbol{R})_{ij}^{2} \right)^{2} \right) \tag{5.14}$$

where ω determines the type of Orthomax (see Table. 5.1). We explore the varimax version [Kaiser, 1958] for optimizing sparsity corresponding to new variables being associated to localized variation modes, i.e, we will have $\omega = 1$.

Setting ω to 1 leads us to reconsider the Equation. (5.14) as the following:

$$\begin{aligned} \boldsymbol{R} &= p \sum_{j=1}^{k} \left(\frac{1}{p} \sum_{i=1}^{p} \left(\boldsymbol{\Theta^2}_{ij} \right)^2 - \frac{1}{p^2} \left(\sum_{i=1}^{p} \boldsymbol{\Theta}_{ij}^2 \right)^2 \right) \\ &= p \sum_{j=1}^{k} \left(\frac{1}{p} \sum_{i=1}^{p} \left(\boldsymbol{\Theta^2}_{ij} - \bar{\boldsymbol{\Theta^2}}_{j} \right)^2 \right) \end{aligned} \tag{5.15}$$

where $\boldsymbol{\Theta} = \Sigma \boldsymbol{R}$, describes the varimax rotated basis, and $\bar{\Theta}_j^2$ the squared mean of the j^{th} column of $\boldsymbol{\Theta}$. Throughout Equation. (5.15), one can note that the investigated varimax rotation can be interpreted as a redistribution of the modes so that each row or

column contains a minimum number of nonzero components, i.e. columns or rows are as sparse as possible.

5.5.2 Modeling Using the Varimax Criterion

After projecting the landmarks into the diffusion wavelet coefficient space we perform PCA, and subsequently rotate the coordinate system to maximize the varimax criterion. The resulting modes represent the variation in the data in a sparse fashion. Let γ_j denote the orthomax eigenvectors such as $\gamma_j = \boldsymbol{R}^{-1}\sigma_j$. Then Equation. (5.12) can be expressed with the orthomax components.

$$\Psi^{i*}_{level} = \left[\Psi_{level}{}^{i*}_{1}\Psi_{level}{}^{i*}_{2}\dots\Psi_{level}{}^{i*}_{K}\right] \tag{5.16}$$

$$V^p_i = \quad \bar{V}^p + \Phi.\left(\bar{\Psi}_{level_j} + \gamma_j\left(\gamma'_j.\Psi^{i*}_{level}\right)\right), \tag{5.17}$$

As it will be shown in the results section (Section. 5.6) the orthomax modes exhibit local variations in most modes, unlike PCA shape modes that are ordered according to variance, and show global variations in the modes.

An overview of the model building process (including the representation component) is given in (Alg.1). The resulting model holds information about the diffusion wavelet tree, the orthomax components, and coefficient variation constraints $\langle \mathbf{\Phi}, \mathbf{R}, \gamma \rangle$.

To summarize; the outcome of this process (see Figure. 5.5) is an efficient shape representation as well as a compact manifold construction with respect to the allowable variations of this representation. We first obtain a topology from the training data and encode it in a diffusion kernel, that defines a diffusion process [Coifman and Lafon, 2006] across the set of landmarks. It can either be based on their distance, or on their mutual dependencies [Langs and Paragios, 2008] observed in the training data. Given this kernel, we build a hierarchical wavelet representation of the shape variation. Finally we build a sparse model of the individual levels of this representation with help of the *orthomax* criterion. Keep in mind that the main goal is to capture meaningful structures based on their behavior in the potentially very small training data set. This representation, along with the manifold can now be used for image based inference in a new example.

It is worth mentioning that the orthomax criterion also allows rotations of subsets of the shape model, while preserving the rest of the modes unvaried. In another context, it is

Algorithm 4 Multi-Scale Representation: Off line Training

Input:

$\boldsymbol{T}$: Diffusion Operator,

$(\mathbf{S}_i)_{\mathbf{i}\in\textbf{Training Shapes}}$: Training landmark deviation from the mean shape

1. Construct the wavelet diffusion tree Φ [Coifman and Maggioni, 2006].
2. Calculate Diffusion Wavelet Coefficient for the training shapes S_i.
3. For each level.
   (a) Compute covariance matrix of every diffusion wavelet coefficient level.
   (b) Calculate a model parameterization (basis, coefficients) based on the orthomax criterion.

Output: Orthomax Eigenspace (modes, eigenvalues) and coefficients of the training examples.

interesting to note that our parametrization results in a separation of sub-parts similar to the n-Cut proposed in [Nain *et al.*, 2007], but replaces the hard splitting by a continuous basis transform.

5.6 Experimental Validation

Throughout this experimental validation section, we evaluate the multiscale shape prior based on a shape reconstruction task. The basic idea is to learn a prior within our training data set. Afterward we project a test volume shape onto the prior so as to estimate how close a projected test shape can be comparing to its ground truth. A common prior of a shape model technique would consist of the mean shape and the eigenvectors of the shape landmarks. For our approach the prior will gather the mean shape, the wavelet tree and the eigenvectors of the diffusion wavelet coefficients from each volume. Later on a reconstruction phase is necessary to compare the modified test shape and the relative supervised segmentation. One of the evaluation indices is indeed the mean squared error between the ground truth and the reconstructed test volume.

We evaluate the algorithm on two medical imaging applications to assess the perfor-

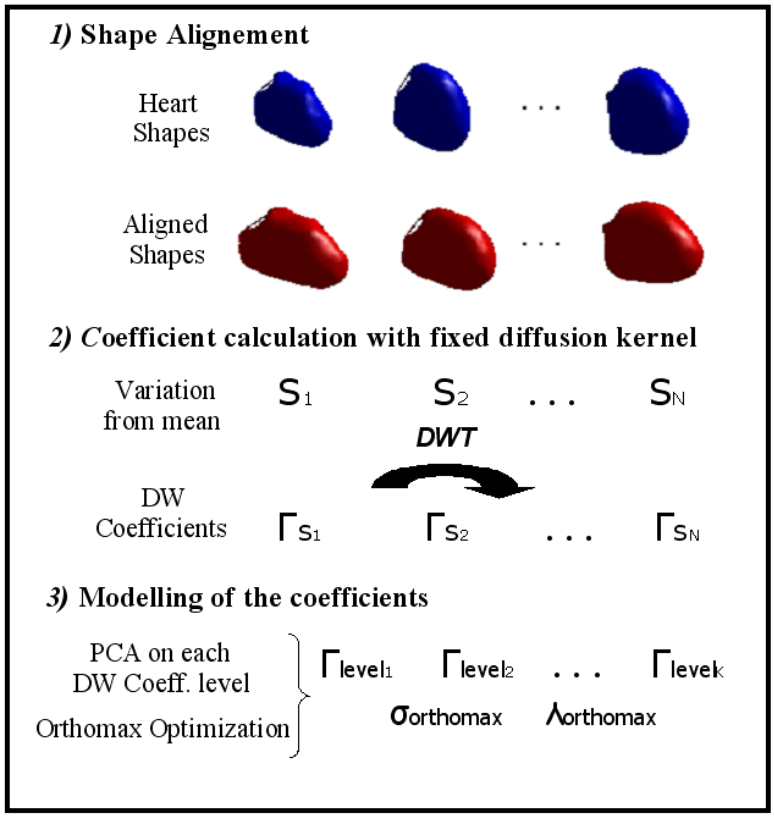


Figure 5.5 — Scheme of Diffusion Wavelet Coefficient Process.

mance of the method in terms of representation, manifold construction and knowledge-based segmentation. The first example is the segmentation of the left ventricle (LV) of the heart using computed tomography images, and the second the segmentation of calf muscles from T1 Magnetic Resonance Images. While in the first case, the performance of the extraction of image support is acceptable, things are far more complicated when considering the muscle images. As described in Section. 3.6, this is due to the fact that for the left ventricle the separation of tissue and lumen is possible while the calf images do not exhibit clear separation between different muscles, and local deformations are far more pronounced in this structure.

To assess our model, experiments were carried out on the following data:

1. 25 MRI calf muscles divided into two groups: 20 healthy control patients and 5 unhealthy cases. For each volume there are 90 slices of 4mm thickness, and

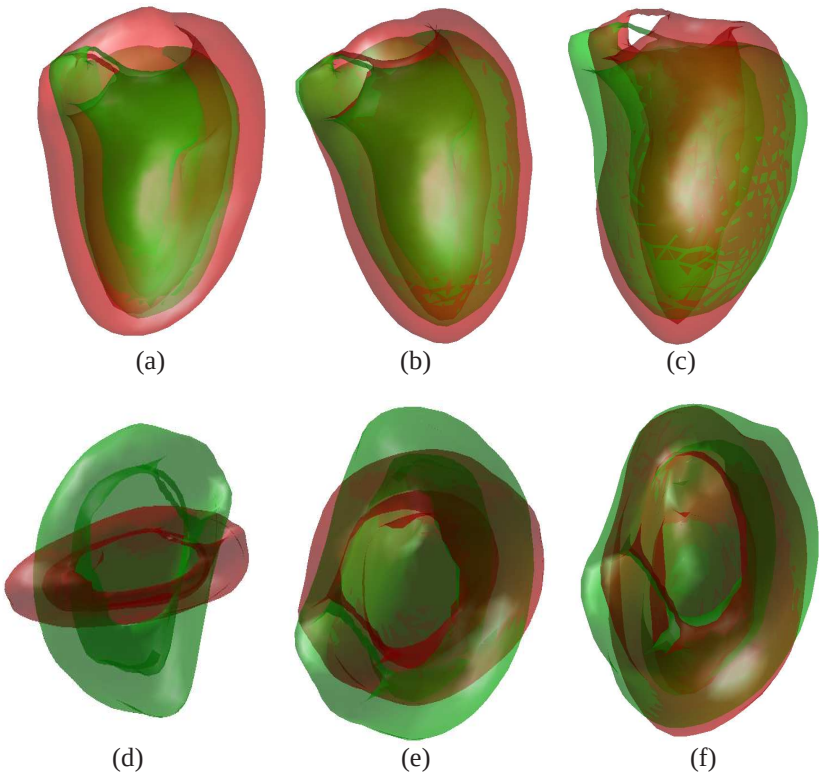


Figure 5.6 — Reconstructed surfaces for heart left ventricle CT data using projected wavelet coefficients on the set of principal components that represent 99% of the total variance at the first level of decomposition. The axial view surfaces placed in the first row represent the $\pm 3sqrt(\lambda_i)$ from left to right. Second row represent the saggital view

with voxel spacing 0.7812x0.7812x4 mm acquired with a 1.5T Siemens scanner. Standard of reference annotation by experts for the Medial Gastrocnemius (MG) muscle, was available (see Figure. 4.4.b). Correspondences for 895 landmarks on the surfaces were obtained by an MDL based optimization [Langs *et al.*, 2007].

2. A set of 25 CT volumes of the heart, with an approximate voxel spacing of 1.5 mm, for which 90 anatomical standard of reference landmarks, and a set of 726 control points for the left ventricle was available, also with available ground truth segmentation from experts Figure. 4.4.a concerning the diastole and the systole.

For each data set we model the shapes of the structures through diffusion wavelet representation in a leave-one-out cross validation strategy. We train the model on 24

	Heart Data	**Calf Data**
Gaussian Model	1.6154	2.1277
DW Model with spatial kernel	0.0755	0.1485
DW Model with Shape Map kernel	0.1100	0.1796

Table 5.2 — Full Landmark Reconstruction Error (in voxel) with regard to three different shape models for heart and calf data sets.

cases, and perform model reconstruction or search on the remaining case not used for modeling. To assess the diffusion modeling approach, we compute two measures: (i) the reconstruction error between the test shape, together with its approximation by the model at different scales, and (ii) the search performance. In the following we will focus on the reconstruction accuracy. The search performance will be evaluated in detail in Chapter. 6. We compare the reconstruction error of Gaussian shape models and the proposed diffusion wavelet model, evaluate two different diffusion wavelet kernels: 1. the spacial proximity of landmarks, and 2. a kernel based on a shape map distance of the landmarks [Langs and Paragios, 2008]. Quantitative results are given by the landmark error between the true shape, and the reconstruction result with one of the three models. Qualitative results are shown in Figure. 5.7.

The main concern is to see how far our model is able to detect the local shape variations based on different kernels. To illustrate the orthomax representation, in the Figure. 5.6 we show the heart reconstructed surfaces using projected wavelet coefficients on the set of principal components at the first level, where the surfaces represent the coefficient values $\pm 3sqrt(\lambda_i)$ from left to right. Comparison of the reconstruction error between the diffusion wavelet model and the reference model is calculated as an average surface error for all test shapes.

In Tab.(5.2), we report errors of a Gaussian reference model, and two diffusion wavelet models. In Figure. 5.7 the reconstruction of the projected shape model (heart/muscle) is depicted, this reconstruction starts from the projection of the diffusion wavelet coefficients in respectively the first and the last level, and then extracting a new subset of coefficients from the eigenvectors that constitutes 99% of the variation in

the correspondant level. To keep the variation in reasonable limits, the shape parameters are also restricted to $\pm 3\sqrt{\lambda_i}$.

Concerning the varimax criterion application, one can easily notice in Figure. 5.8 that while the PCA modes demonstrate several spatially distributed effects within each mode, the varimax modes in the other hand show nicely isolated effects. Moreover in Figure. 5.9, we show the 'flattening' of the eigenvalue spectrum carried out by the varimax rotation where the respective modes as well as variances are plotted. This simple, yet powerful modification of PCA enables us to optimize sparsity leading to localized modes of variation, which is more suitable for applications with sparse parameterizations like the often local pathological variations we are focusing on.

5.7 Contributions

We described in the current chapter a novel approach to represent prior knowledge for image segmentation using diffusion wavelets. During the learning stage, the underlying diffusion operator is learned from the data. The corresponding diffusion wavelet structure is build, and the wavelet coefficients for the training shapes are calculated. The diffusion kernel incorporates a notion of soft connectivity between landmarks by encoding inter-object relationships. Consequently the approach is able to provide a hierarchical representation. It models both geometry and appearance, and can learn arbitrary topologies, encoded in the kernel. Results of the segmentation of MRI and CT data show that the method has promising reconstruction performance, and is able to model shape variation with higher accuracy than a standard Gaussian model. Until nows the diffusion wavelet have largely been confined to the signal processing field, and have not yet been exploited for medical image segmentation.

The learning of the wavelet domain topology from the training data enables the algorithm to represent complex structures like groups of muscles in a single model. The way of learning this domain by either using local neighborhoods, or deformation complexity makes an adaptation to the data possible. Its impact on different data is subject of future research.

The current chapter introduces the following ideas;

- A novel segmentation framework based on diffusion wavelet models.

- The learning of the model parameterization: we learn the wavelet domain, instead of using a pre-defined manifold.
- The approach encodes hierarchies and soft connectivity properties by means of the diffusion kernel.
- The method provides the ability for search paradigm based on local appearance features.
- The computational efficiency in high dimensional spaces regarding the orthomax technique.

Undoubtly the most attractive characteristic of the wavelet kernel modeling approach is the ability to represent local shape differences in an effective manner. Moreover the application of diffusion wavelets to segmentation is quite novel. We addresses an important problem, namely replacing the fixed shape representations as in ASM/AAM by multi-scale priors which reflect inter-dependencies in training data.

As a comparaison with the work of [Nain *et al.*, 2007], and with respect to segmentation, spherical wavelets represent a sub-case of diffusion wavelets confined to a spherical manifold. With regard to registration approaches, the prime difference with our scope is that we are building a generative model of shape variation, and use it for segmentation while at the same time using the orthomax criterion to obtain an optimal subdivision of the shape parameterization.

Furthermore we have investigated the orthomax criterion for principal components rotation. This computationally simple technique enables us to optimize sparsity and to have more localized modes of variation. Such an approach is of a great interest regarding pathological changes in the anatomical structures. The combination of both diffusion wavelet and varimax rotation is a promising way to model complex shapes with locally different variability.

In the next chapter we demonstrate how we can perform search in new data. We use the diffusion wavelet shape prior introduced in this chapter, and a local appearance representation.

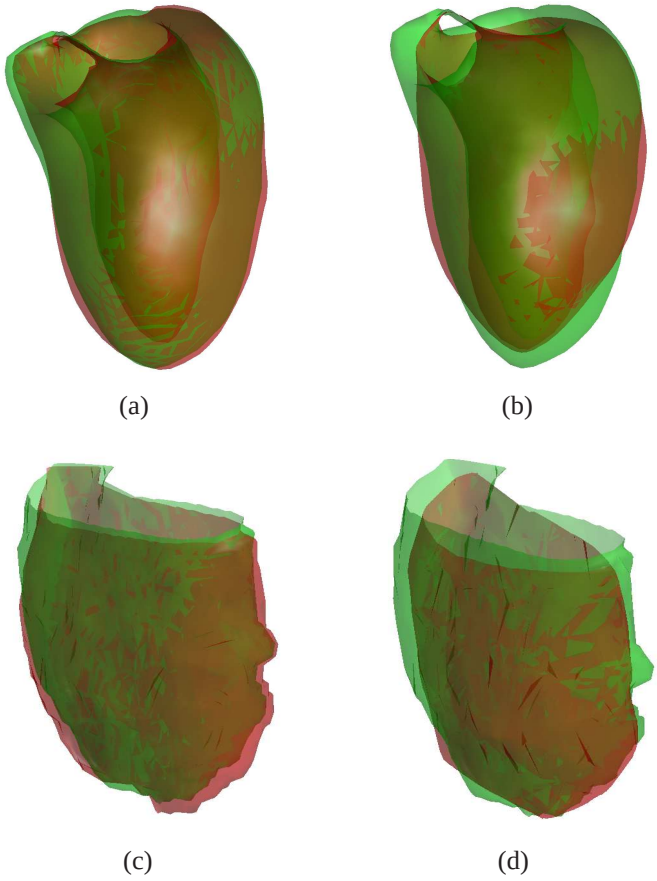
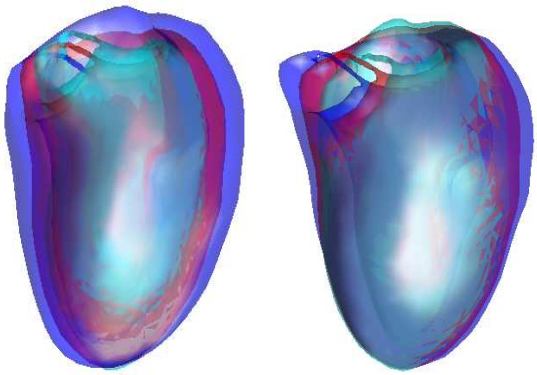


Figure 5.7 — Diffusion Wavelets Model Reconstruction. First row: Heart results and second: Calf muscle. Data, green: standard of reference segmentation, red: reconstruction result for a. finest scale and b. coarsest wavelet scale.



(a) (b)

Figure 5.8 — Data reconstruction through (a) global PCA and (b) localized orthomax rotated modes of the DW coefficients models. The surfaces representing the $\pm 3sqrt(\lambda_i)$ are respectively colored in cyan, red and blue. Note the local deformation captured by the orthomax mode.

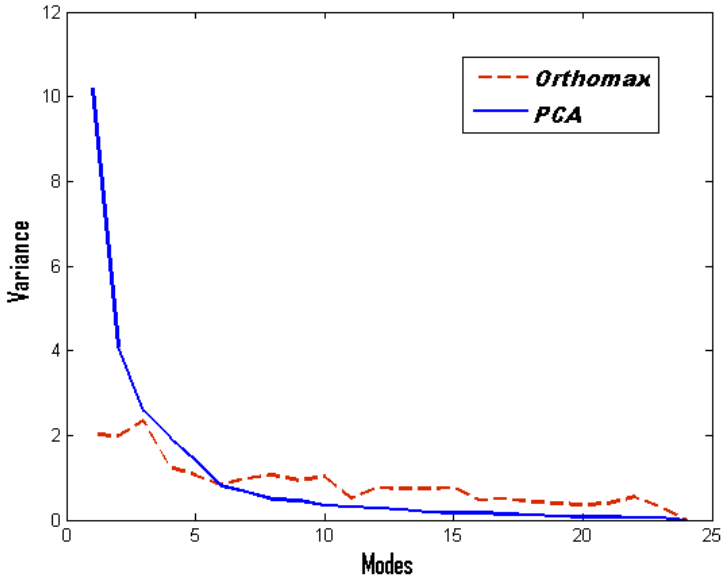


Figure 5.9 — Comparaison between PCA and Orthomax DW eigenvalues

CHAPTER 6

Search Algorithms Performance

"Anyone who has never made a mistake has never tried anything new."

Albert Einstein

6.1 Introduction

In this chapter we explore approaches to perform search with a model that is build as described in the previous chapters. The following sections focus on the inference of landmark positions from image data, and the subsequent estimation of the shape model parameters. We will first review existing work, and will then give a brief outline of the two prime concepts we use: 1. Canonical Correlation Analysis (CCA) to learn the relation between observed appearance and optimal model parameter update; 2. Representing appearance by means of local descriptors.

After this, we will introduce a classifier based local appearance representation, and will explain how to learn it during training, and use it in conjunction with the proposed shape models during search. The performance and experimental validation are finally reported in Section. 6.6.

6.2 Related Work

A large amount of research has been published regarding appearance search. Active Appearance Model (AAM) were proposed in [Cootes *et al.*, 1998b]. An approach that integrates model search with classification approaches was proposed in [van Ginneken *et al.*, 2002], where optimal feature classification is proceeded in order to select the best set of features corresponding to each landmark, a *k*-Nearest Neighbors *(*kNN) classifier is then used throughout the search to find optimal displacements over the landmarks.

[Zhan and Shen, 2006] proposed a gabor filter based SVM (Support Vector Machine) framework in order to extract features from 3D ultrasound prostate data and estimate the likelihood of a voxel to lie in a prostate tissue.

[De Bruijne and Nielsen, 2004] introduced a shape model inference on the basis of pixels classification using particle filter. They preserve the global shape and appearance from the deformable techniques, however get rid from the problems of localized appearance variation and initialization issue by a maximum likelihood shape inference on pixel classification. The authors were indeed inspired by object tracking literature.

[Mitchell *et al.*, 2002] were the pioneer to introduce a 3D model exploiting AAM using volumetric texture to represent appearance. Other related approaches combining local features with standard shape models were developed by [Scott *et al.*, 2003] [Qian

et al., 2005] [Zheng *et al.*, 2007].

In the following we will explain the basic search method employed by AAMs. In the next section we will contrast it with the approach based on classifier driven appearance representation.

6.2.1 Active Appearance Model

Active Appearance Model Search [Cootes *et al.*, 1998b] represent the most well known region based feature appearance model. AAMs, similar to ASMs algorithm (see Section. 3.4), require a set of corresponding landmarks in a set of training examples. AAMs then combine the shape and appearance in a single model. This generative model is able to synthesize realistic images of the modeled data. Similarly to ASMs, pre-processing steps including Procrustes analysis and PCA are required for the shape $\mathbf{x}$, and analogously for the pixel intensities $\mathbf{g}$ that are warped into a shape-free reference frame and sampled. A new instance for the appearance is then produced as a linear statistical system of the appearance model components and subsequently deformed according to the shape model components,

$$\mathbf{g} = \bar{\mathbf{g}} + \Phi_g \mathbf{b}_g \tag{6.1}$$

where $\bar{\mathbf{g}}$ is the mean normalized appearance vector, Φ_g is a set of orthogonal independent eigenvectors of gray value model and finally $\mathbf{b}_g$ is a set on gray level parameters. A combined appearance and shape model is generated by a concatenated parameter vector $\mathbf{b}$:

$$\mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \begin{pmatrix} \mathbf{\Phi}_{c,s} \\ \mathbf{\Phi}_{c,g} \end{pmatrix} \tag{6.2}$$

That holds information regarding both the appearance and the shape. Here $\mathbf{W}_s$ represents a diagonal matrix of weights assigned for every shape parameter, resulting in the following expression:

$$\mathbf{b} = \mathbf{\Phi}_c \mathbf{c} \tag{6.3}$$

where $\mathbf{c}$ and $\mathbf{\Phi}_c$ corresponds to the appearance vector parameter and to the eigenvectors controlling both the shape and the gray levels. A synthetic image for a given parameter $\mathbf{c}$ is produced by the appearance vector $\mathbf{g}$ and warping it with the shape $\mathbf{x}$ as in Equation. (6.4).

$$\begin{aligned} \mathbf{x} &= \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{c} \\ \mathbf{g} &= \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{c} \end{aligned} \tag{6.4}$$

with

$$\begin{aligned} \mathbf{Q}_s &= \Phi_s \mathbf{W}_s^{-1} \Phi_{c,s} \\ \mathbf{Q}_g &= \Phi_g \Phi_{c,g} \end{aligned} \tag{6.5}$$

Through the use of iterative updating scheme, the model parameter $\mathbf{c}$ can be fitted rapidly to a new volume image exploiting the $L2$-norm as a cost function.

A more rapid search method was proposed in [Cootes and Taylor, 2001b], where the dependence between model parameter update and current residual appearance error is learned during the training phase.

The AAM model vector parameter $\mathbf{p}$ defines the generated synthetic image. During each step of the training the model texture and the sampled image patch respectively $\mathbf{g}_m(\mathbf{p})$ and $\mathbf{g}_s(\mathbf{p})$, the residual vector $\mathbf{r}$ is then parametrized by $\mathbf{p}$ and computed like in Equation. (6.6).

$$\mathbf{r}(\mathbf{p}) = \mathbf{g_s} - \mathbf{g_m} \tag{6.6}$$

Using the first order Taylor expansion of $\mathbf{r}$, one obtains such approximation:

$$\mathbf{r}(\mathbf{p} + \delta\mathbf{p}) \approx \mathbf{r}(\mathbf{p}) + \frac{\partial \mathbf{r}}{\partial \mathbf{p}} \delta p \tag{6.7}$$

The key idea behind the search will be to consider it as an optimization problem, where the goal is to minimize the difference between a new image and the synthesized one, the regression will be then expressed through the parameter update as:

$$\delta \mathbf{p} = -\mathbf{R}\mathbf{r}(\mathbf{p}) \tag{6.8}$$

The calculation of the derivative matrix $\mathbf{R}$ determines the optimization efficiency. The computation is either relying on a multivariate linear regression [Cootes *et al.*, 1998b], or alternatively on a numeric differentiation [Cootes and Kittipanya-ngam, 2002]. Further details, notably the choice of shape parameters weights, and how to optimize the model fitting regarding the derivative matrix $\mathbf{R}$ can be found in [Cootes and Taylor, 2001a].

In the review of [Cootes and Kittipanya-ngam, 2002], various approaches have successfully improved AAM search performance, by modifying the original algorithm scheme. A few examples are:

- ShapeAAMs [Cootes and Kittipanya-ngam, 2002] update only shape and pose parameters during search, while texture parameters are directly calculated from the training examples by using projections of image textures.

- DAMs: Direct Appearance Model [Hou *et al.*, 2001] predict shape parameter directly from texture information, and thus without combining from shape and texture as in AAM. The DAMs result indeed in improving convergence and solution accuracy.

- Fast AAM using CCA [Donner *et al.*, 2006] where the authors exploit the canonical correlation analysis for a faster search process, providing this way a more precise estimation for parameter update comparing to standard model. Another merit of this method is the fewer number of examples necessary for building the training set, which has an important impact on accelerating the model construction. A direct consequence is the optimization of the regression matrix.

As for local texture descriptors, several operators have been proposed in the literature in order to extract texture features, to cite but a few; steerable filters, SIFT features, moment invariants and shape context. An exhaustive review and evaluation concerning these point descriptors is established by [Mikolajczyk and Schmid, 2005]. It is indeed worth mentioning that steerable filters along with moments show highly effective performance in comparison with other descriptors, particularly thanks to their low dimensionality and reliability.

6.2.2 Canonical Correlation Analysis (*CCA*)

As far as linear and orthonormal transformations are concerned, PCA represents the optimal technique as it minimizes the reconstruction error between the original data and the reconstructed one, in the mean square sense. Nevertheless, apart from PCA there are various other linear methods in the literature that are considered as more appropriate to regression exercise. Let's cite for example Multivariate Linear Regression (MLR), Partial Least Squares (PLS), or yet Canonical correlation analysis (CCA) [Melzer *et al.*, 2003].

Canonical correlation analysis (CCA) [Hotelling, 1936] is a robust method of measuring the linear relationship between two multidimensional data sets. It is a powerful tool able to find out the optimal bases with respect to correlations corresponding to each set of variables. One important characteristic is obviously the invariance regarding any similarity transformation. Originally this statistical tool has been more exploited in domains like economics, meteorology, and only recently in medical imaging and computer vision [Melzer *et al.*, 2003] [Donner *et al.*, 2006].

CCA shares some similarities with PCA, mainly the goal and application of data reduction, on the other hand however CCA is more suited for regression as it takes into account the correlation between the two sets of measurements.

More practically let's consider a random variable $\mathbf{x} \in \mathbb{R}^p$ with zero-mean, similarly $\mathbf{y} \in \mathbb{R}^q$, and the following linear combinations $x = \mathbf{x}\mathbf{w}_x$ and $y = \mathbf{y}\mathbf{w}_y$. Maximizing the function ρ (see Equation. (6.9)) with respect to $\mathbf{w}_x$ and $\mathbf{w}_y$ is equivalent to finding the maximum canonical correlation.

$$\begin{aligned} \rho &= \frac{E[xy]}{\sqrt{E[x^2]E[y^2]}} = \frac{E[\mathbf{w}_x^T\mathbf{x}\mathbf{y}^T\mathbf{w}_y]}{\sqrt{E[\mathbf{w}_x^T\mathbf{x}\mathbf{x}^T\mathbf{w}_x]E[\mathbf{w}_y^T\mathbf{y}\mathbf{y}^T\mathbf{w}_y]}} \\ &= \frac{\mathbf{w}_x^T\mathbf{C}_{xy}\mathbf{w}_y}{\sqrt{\mathbf{w}_x^T\mathbf{C}_{xx}\mathbf{w}_x\mathbf{w}_y^T\mathbf{C}_{yy}\mathbf{w}_y}} \end{aligned} \tag{6.9}$$

Indeed we call out canonical variates the projections onto $\mathbf{w}_x$ and $\mathbf{w}_y$, i.e. x and y. The total covariance matrix, showed in Equation. (6.10), represents a block of matrix composed by $\mathbf{C}_{xx} \in \mathbb{R}^{p*p}$ and $\mathbf{C}_{yy} \in \mathbb{R}^{q*q}$, where the latter are the within set covariance matrix of respectively $\mathbf{x}$ and $\mathbf{y}$. On the other hand we denote by $\mathbf{C}_{xy} \in \mathbb{R}^{p*q}$ the between

set covariance matrix.

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & \mathbf{C}_{yy} \end{bmatrix} = E\left[\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}^T \right] \tag{6.10}$$

The canonical factors can thus be obtained by a Singular Value Decomposition (SVD) of the matrix $\mathbf{F}$

$$\mathbf{F} = \mathbf{C}_{xx}^{-\frac{1}{2}} \mathbf{C}_{xy} \mathbf{C}_{yy}^{-\frac{1}{2}} \tag{6.11}$$

Let the SVD of $\mathbf{F}$ be expressed as $\mathbf{F} = \mathbf{U}\mathbf{D}\mathbf{V}^T$, we can then extract the i^{th} canonical factor as following

$$\begin{cases} w_x^i = \mathbf{C}_{xx}^{-\frac{1}{2}} \mathbf{u}_i \\ w_y^i = \mathbf{C}_{yy}^{-\frac{1}{2}} \mathbf{v}_i \end{cases} \tag{6.12}$$

In Equation. (6.12) one finds $\mathbf{u}_i$ and $\mathbf{v}_i$ that describe respectively the i^{th} column of the matrices $\mathbf{U}$ and $\mathbf{V}$, whereas the corresponding canonical correlations are represented by the di- agonal components of $\mathbf{D}$.

Comparing to other correlation techniques, CCA is dependent on the coordinate system in which the variables are described. CCA has also some very attractive properties over Multivariate Linear Regression mainly improving predictive accuracy and scale invariance. In contrast to PCA, CCA is simplifying the obtained pose estimation on the manifold.

6.2.3 Feature Space and Geometric Descriptors

Curvature Estimation Knowledge of surface curvature is of particular importance to a wide range of applications such as pattern matching, computer graphics, or classification. We opt for the solution proposed by [Rieger *et al.*, 2004] which does not only offer a robust curvature estimation for 3D surface, but also has the merit to escape from the majority of problems usually inherited by common estimation techniques. Indeed the proposed estimator works on the orientation field of the surface. The Gradient Structure

Tensor (GST) will be exploited in order to obtain the orientation field and the description of local structure.

The curvature k at a point p in a tangent direction $\mathbf{t}$ on a surface is defined as;

$$k_t(p) = \|\nabla_t \mathbf{N}\| \tag{6.13}$$

To obtain the the principal curvatures and shape descriptor, there are two crucial steps:

1. Determine the vector field normal $\mathbf{N}$ and the principal directions t_1 and t_2, for which the curvatures are extremal.

2. Resolve the discontinuity problem of $\mathbf{N}$ and estimate $\|\nabla_t \mathbf{N}\|$.

The robust estimation is primarily established through the generic tool gradient structure tensor $\bar{G}$ able to analyze local image structure. One can define the GST as, $\bar{G} = v\bar{v}^t$, where $v = \nabla I$ and I represent the gray value image. A first processing level is an eigenvalue decomposition of $\bar{G}$, leading to ordered eigenvalue and corresponding eigenvectors $\{v_i\}$. The first eigenvector is then aligned with the normal surface N, whereas the two following ones are relied to the principal surface directions as shown in Equation. (6.14).

$$v_1 \longleftrightarrow \mathbf{N} \quad , \quad v_{2,3} \longleftrightarrow \mathbf{t}_{1,2} \tag{6.14}$$

The following step consists on resolving the discontinuity problem, in other words calculating the principal curvatures by mapping v_1 into a continuous representation fulfilling the current statement $\|\delta M(v)\| = K\,\|\delta v\|$. In fact the authors propose a solution depicted by Equation. (6.15). More computational details can be found in [Rieger *et al.*, 2004] and references therein.

$$|k_{1,2}| = \frac{1}{\sqrt{2}} \left\|\nabla_{v_{2,3}} M(v_1)\right\| \tag{6.15}$$

Gabor Jets Gabor Jets are the outputs from a set of Gabor filters [Movellan, 2002]. They are able to exploit salient visual properties such as spatial localization, orientation selectivity, and spatial frequency properties. Moreover, one of the advantages of Gabor

Jets is that their phase is invariant to edge positions, if the filter responses were converted to amplitude and phase. These filters are indeed having a great success in faces recognition domain.

A Gabor filter is represented by two functions, a complex sinusoidal $s(x, y)$ called *carrier* and a Gaussian-shaped function known as the *envelop*;

$$g(x, y) = s(x, y)w_r(x, y) \tag{6.16}$$

The complex sinusoidal carrier function is defined as following

$$s(x, y) = exp\Big(j\Big(2\pi(u_0x + v_0y) + P\Big)\Big) \tag{6.17}$$

where u_0 and v_0 define the spatial frequency in Cartesian coordinates and P the phase. For the Gaussian envelop it can be written as;

$$w_r = Kexp\Big(- \pi\Big(a^2(x - x_0)^2_{r(\theta)} + b^2(y - y_0)^2_{r(\theta)}\Big)\Big) \tag{6.18}$$

where K is a scaling factor, (x_0, y_0) the spatial coordinates of the Gaussian envelop peak, a and b the scaling parameters for the two axis of the Gaussian, and finally $r(\theta)$ stands for a rotation operation describing a clockwise rotation with the angle θ.

By utilizing the Gabor filters we aim to extract the image features for each of the defined patch window. The amount of extracted features will be mainly related to the fixed number of angles rotation and frequencies.

6.2.4 Classification Approaches

In Section. 6.4we will rely on classifiers for the representation of texture. In the following we give a brief review of three classifiers, and their characteristics relevant for our task. Let's first review some powerful state of the art weak learning techniques, let's recall that a weak learning algorithm is a processing that can always generate a hypothesis with a weak edge for any distribution.

Support Vector Machine The Support Vector Machines (SVMs) algorithm [Vapnik, 2000] estimates the optimal classifier in the original space while at the same time supervising its complexity by using these tighter predictors of the generalization power of the model. Furthermore to the theoretical arguments for its asymptotic optimality, Support Vector learning has been empirically proven to be robust to over-fitting and to well generalize even in case of small data sets.

AdaBoost [Freund and Schapire, 1996] as the SVM technique, boosting method belongs to the family of supervised and discriminative classification used in the context of medical segmentation, and Adaboost more particularly shows better performance than conventional boosting methods [Viola and Jones, 2001]. Indeed it can be considered as a constraint gradient descent in an error function with respect to the margin. The approach has even been extended in a cascade of Adaboost still by Viola and Jones resulting in a faster detection that is for sure helpful for large data sets.

Decision Forest Generally speaking a random decision forest represent a collection of deterministic decision trees. Mostly applied in the machine learning field, it has also been proved to be useful and successful within the medical images [Criminisi *et al.*, 2009]. Combined all along with learned visual features, decision forest can be constructed in such a way to detect and capture anatomical structures, leading to high computational performance.

6.3 Image-based Search

Following the initial search approach described in Chapter. 4 we can perform search based on the local appearance in the volume. The search with the diffusion model representation, and appearance patch models $(P_i)_{i=1,\ldots,N}$ for each landmark is performed in an iterative manner, starting from a coarse initialization obtained by, e.g., atlas registration. The appearance model is based on a local texture patch model at the landmark positions. Similar to a standard shape model inference scheme, the landmarks positions in new data are estimated by an energy minimization involving both shape prior and appearance costs:

1. The landmark positions of the test volume $\hat{\mathbf{V}}$ are updated according to a local

appearance model. For each landmark the position with highest probability of the corresponding local texture patch being consistent with regard to the learn texture model P_i is chosen within a neighborhood to of the current position estimate. We consider $P(\mathbf{x}_i^j)$ as the learned texture patch for the correct landmark position $\mathbf{x}_i$ in the initial training volume. During search for each landmark position we look for a better appearance fit in a local neighborhood $\mathcal{N}$. That is we use correlation as the similarity measure, and we have $C_i^j(\mathbf{x})$ as the correlation between the current candidate patch $p(\mathbf{x})$ and the model patch $P(\mathbf{x}_i^j)$ normalized within the neighborhood, i.e. $\int_{\mathbf{x}\in\mathcal{N}} C_i^j(\mathbf{x}) = 1$, then the image support is

$$\xi_i = mean_{j=1,\ldots,n}\left(\frac{C_i^j(\mathbf{x})}{\int_{\mathbf{x}\in\mathcal{N}\setminus\mathbf{x}_i^j} C_i^j(\mathbf{x})}\right). \tag{6.19}$$

The image support is thus computed for every landmark in $\mathcal{V}$ from the local appearance behavior at the corresponding positions in the training shape.

2. After fitting the shape to the image data, its variation is constrained by the diffusion wavelet shape variation model. The landmarks are projected into the orthomax coefficient space as described in Sec 5.5.1. The constraints learned during training are applied, and based on the resulting parameter values the shape is reconstructed.

This procedure is iterated while during each iteration, the corresponding $\mathbf{V}'$ is reconstructed to re-estimate the shape. After convergence the final reconstruction $\mathbf{V}'$ is an estimate of the true shape inferred from the data, and the prior model.

6.4 Appearance Classifiers Search Scheme

6.4.1 Learning a Classifier for Appearance Modeling

An alternatively to representing appearance by local features, or just the mean appearance, is to use classifiers to obtain a more flexible and adaptive representation.

For each landmark we can train a classifier, that is able to differentiate between background and the landmark. Likely there is a very high level of ambiguity resulting from such a classification if it is applied to the entire volume. However, we can treat the positive labels, or probabilistic weights resulting from the classification as hypotheses. In a second step the shape model is used to pick the most likely hypothesis from this set.

We adopt recent developments in machine learning that explores the use of weak classifiers and arbitrary image features. For the heart muscle, our feature space involves the (i) gradient phase and magnitude, (ii) structure tensor plus their (iii) curvature [Rieger *et al.*, 2004] and (v) the responses to Gabor filters with different phases and orientations.

Once feature vectors have been extracted from the input shapes, we can estimate the optimal classifier as in a traditional machine learning framework. Several options can be employed, among them one can cite Support Vector Machines (SVM) that represent a powerful clustering technique searching for a hyper-plane and a normal vector with the least possible norm, leading to an effective separation between the labeled data. Boosting methods are weak linear classifiers capable of generating outstanding classification results upon proper integration. Another interesting work recently presented is compressed sensing [Donoho, 2006] where the main idea consists of recovering from a set of subspaces the least possible number of examples able to express the observations under a sparsity assumption. These methods were more explained in details in Section. 6.2.4, and Boosting algorithm will be tested over our heart anatomical data set in Section.6.2.4.

We use the Adaboost classifier, as it is provably effective and suits better to the characteristics of our data and feature space in which we expect substantial variability across training subjects, and anatomical sites. Starting from our feature space, we apply Gentle Adaboost [Friedman *et al.*, 2000] to obtain a local appearance prior for the search in new data. The boosting process aims at building a strong classifier by combining a number of weak classifiers, which need only be better than chance. For this we call upon a sequential learning process: at each iteration, we add a weak classifier. It is the basic learning algorithm introduces by [Viola and Jones, 2001]. For each landmark the classification problem is as a two class training set (background vs. landmark) that can be represented as: $S = \{(x_i, y_i)\}_{i=1}^{l} \subset \mathbb{R}^N \times \{-1, +1\}$. Let's denote $D_t(i)$ as the weight of the distribution on training example i on round t. Initially all weights are set equally $D_1(i) = 1/l$

We use these classifiers to locate landmarks during the segmentation process. The classifiers detect landmarks present on the ventricle muscle wall against background. This is a very different strategy in comparison to standard search methods [Cootes *et al.*, 1995]. The main search strategy is: extract features from the volume, for each landmark obtain a few candidate positions with a very strong classifier response, fit the DW model to these candidates, and determine the candidate configuration with the

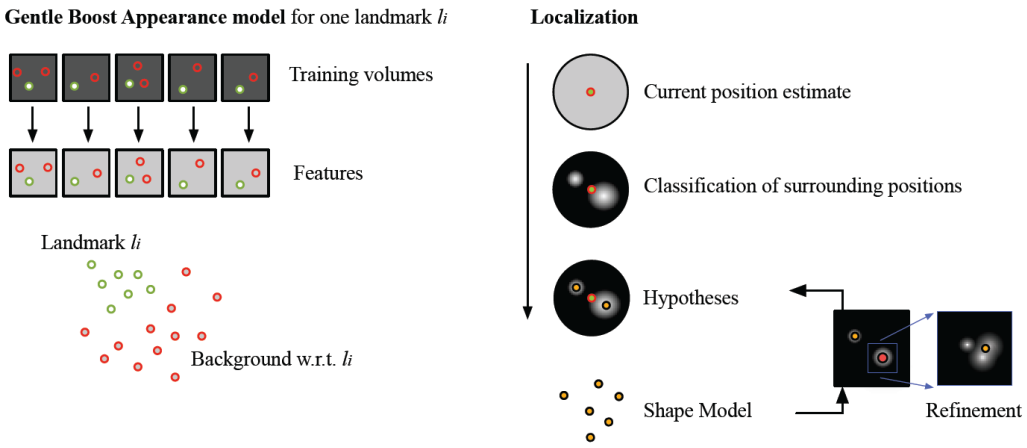


Figure 6.1 — Scheme of the appearance model: Based on local features, and a classifier we can assign each position in the volume an evidence value for landmarks presence. This results in a set of hypotheses for landmark positions, that are verified by the shape model constraint.

highest plausibility with regard to the shape prior. After continue with the local search at the current landmark estimates constraint by the diffusion wavelet model. During the shape model fitting we check which candidates have the highest plausibility with the trained diffusion wavelet model.

In Figure. 6.1 the scheme of the model search is depicted. For each landmark the search volume $\mathbf{V}$ is projected into a hypothesis space $\mathbf{V}_i^H$ that reflects the evidence for the landmark presence for each point in the volume. This results in a position hypothesis $\hat{\mathbf{x}}_i$ for each landmark. The set of landmark hypotheses $\langle \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_n \rangle$ is tested with the diffusion wavelet shape model, resulting in a position prediction for each landmark. These predictions are used to generate new hypotheses based on the local image support $\mathbf{V}_i^{H'}$ and the shape model. The hypothesis space is the classifier response on each position in the volume. During the progressing search we just consider the neighborhood of the current landmark location estimate during the last iteration.

The method computes a local feature vector for every voxel and maps it via a GentleBoost classifier to a probability that the voxel belongs to a specific landmark in the object. The classifier is trained from the training data set segmentations. The probabilistic output is constrained by the shape model. The mapping onto the diffusion wavelet coefficient space ensures valid results with regard to the training data. The result of this procedure is a probability for each voxel regarding its match to the structure to be segmented, conditioned on both local and global information.

Therefore, during the search process, the model is used together with the Gentle-Boost classifier trained on the local appearance of the individual landmarks describing the anatomical structure. The hierarchical diffusion wavelet shape model is then fitted to new data based on local appearance captured by the classifier.

6.4.2 The Learning and Localization Algorithm

Let us summarize the learning and search concepts introduced in this section. The method consists of a training phase and a search step. First, the shape model and parameterization, and the local classifiers for the appearance representation are learned. During search they are used to locate and segment structures in new image data.

Learning: During the training both geometry and appearance of the structure of interest are learned.

- Given n examples of the structure of interest location and the corresponding images, we represent the shape variability by a diffusion wavelet shape model as described in Chapter. 5.
- Using the same examples, we extract local features for each training volume at different scales. For each landmark, at each resolution, we construct a set of training samples containing local features and corresponding labels which indicate if the position is the landmark location or the background. Background voxels are chosen randomly in the volume except the particular landmark positions for training.
- A cascade of classifiers that differentiates between background and landmark on increasingly fine scales, and within more constrained neighborhoods is learned. After the first level, we train the classifier only in a neighborhood of the landmark position. This results in higher specificity within the vicinity of the landmark position. To train for the fine local differentiation we take into consideration only the neighborhood of every landmark candidate in each training image. We train a classifier for each landmark and retain only the ones with solid performance or wide-margins between the different classes.

Segmentation: Using both diffusion wavelet shape prior and appearance priors, we perform the structure localization as follows: the process is initiated with the mean shape, and proceeds in an iterative manner,

- Perform a local search for the most probable landmark positions using the trained classifiers,
- Constrain the solution using the diffusion wavelet coefficient constraints, and repeat the previous search steps until convergence.

This results in landmark location estimates in the search image, that are based on the appearance, and the shape constrained learned during the training phase.

6.5 Local Appearance Features CCA Search Scheme

6.5.1 Appearance Representation Using Local Features

We employ a search framework related to Active Feature Models (AFM) proposed in [Langs *et al.*, 2006], where the authors introduce a local descriptor based on steerable filters. The approach is related to Active Appearance Models (AAM) [Cootes *et al.*, 1998a], with the main difference that AFMs describe appearance by means of local features, and infers model updates during search by means of canonical correlation analysis (CCA), which has advantages given noisy data. CCA-AAMs are closely related, but use the full appearance representation similar to AAMs. The benefit of the method is considerable in the presence of complex data, like muscle data, where large parts of the variation within the muscles have low relevance for landmark localization and a small training set has to compromise a decent representation of the texture in the model.

We utilize a Gabor jet with frequencies $\{0.3, 0.6, 0.9\}$ and directions $\{0, \pi/4, \pi/2, 3\pi/4\}$ to describe the local texture at the landmark positions. During training we learn the relation between landmark displacements and the corresponding texture feature change by CCA: model parameters are perturbed randomly generating a large number of displaced model instances. A functional relation is then learned from the resulting feature vectors describing local texture $\mathbf{G}$ and the corresponding model parameter displacement by Canonical Correlation Analysis (CCA) $\delta\mathbf{p}$. We generate a set of synthetic images by perturbing the optimal parameter vector, i.e., $\mathbf{r}\left(\mathbf{p}_{opt} + \delta\mathbf{p}\right)$. The vector $\mathbf{p}_{opt}$ is computed by mapping the training image texture and shape into the model eigenspace, where $\delta\mathbf{p}$ elements are randomly drawn from uniform distributions in the interval $[-1, 1]$ standard deviation. Consequently we obtain m feature vectors with m corresponding parameter displacement vectors.

Given $\mathbf{P} \in \mathbb{R}^{q \times m}$ the set of random displacement vectors, and $\mathbf{G} \in \mathbb{R}^{p \times m}$ the series of relative feature vectors, we are able to execute CCA and compute the canonical factors of these signals. Through this factor analysis we will obtain the following linear combinations $\mathbf{W}_g = (\mathbf{w}_g^1, \ldots, \mathbf{w}_g^{k^*})$ and $\mathbf{W}_p = (\mathbf{w}_p^1, \ldots, \mathbf{w}_p^{k^*})$, respectively, where $i = 1 \ldots k^* \leq k$.

Afterward regression is applied on the leading canonical projections $\mathbf{G}_{proj} = \mathbf{W}_g^T \mathbf{G}$ as well as $\mathbf{P}$. These projections are then used to compute the $p \times k^*$ transformation matrix $\mathbf{l} = \mathbf{P}\mathbf{G}_{proj}^{\dagger}$, where $\mathbf{G}_{proj}^{\dagger} = (\mathbf{G}_{proj}^T \mathbf{G}_{proj})^{-1} \mathbf{G}_{proj}^T$.

6.5.2 Diffusion Wavelet Shape Model Search by Canonical Correlation Analysis

During search, local features are extracted at the current landmark position estimates. Based on the relation learned by CCA, according diffusion wavelet shape model parameter updates are performed . This results in an iterative search approach, that converges to the landmark positions, based on the local appearance, and constraint by the diffusion wavelet shape model.

At each iteration the new prediction for a model parameter update is generated at each iteration, as a substitute of the one calculated in Equation. (6.8). Indeed the new prediction $\delta\mathbf{p}_{predicted}$ can be obtained as $\delta\mathbf{p}_{predicted} = \mathbf{l}\mathbf{r}_{proj}$ where $\mathbf{r}_{proj} = \mathbf{W}_g^T \mathbf{r}_{current}$. As $\mathbf{R}_{cca} = \mathbf{l}\mathbf{W}_g^T$ can be pre-computed during training the final formulation of the prediction function as following

$$\delta\mathbf{p}_{predicted}(\mathbf{r}_{current}) = \mathbf{R}_{cca}\mathbf{r}_{current}, \tag{6.20}$$

Our segmentation framework is summarized as shown in Fig. 6.3

6.6 Experimental Results

To assess the segmentation accuracy, we combined modeling methods detailed in the previous chapters and the listed search methods. Model searches are initialized by a displacement equal to the mean configuration $\pm 10\%$ of the mean volume; We carried a leave one out strategy in order to investigate the association behavior.

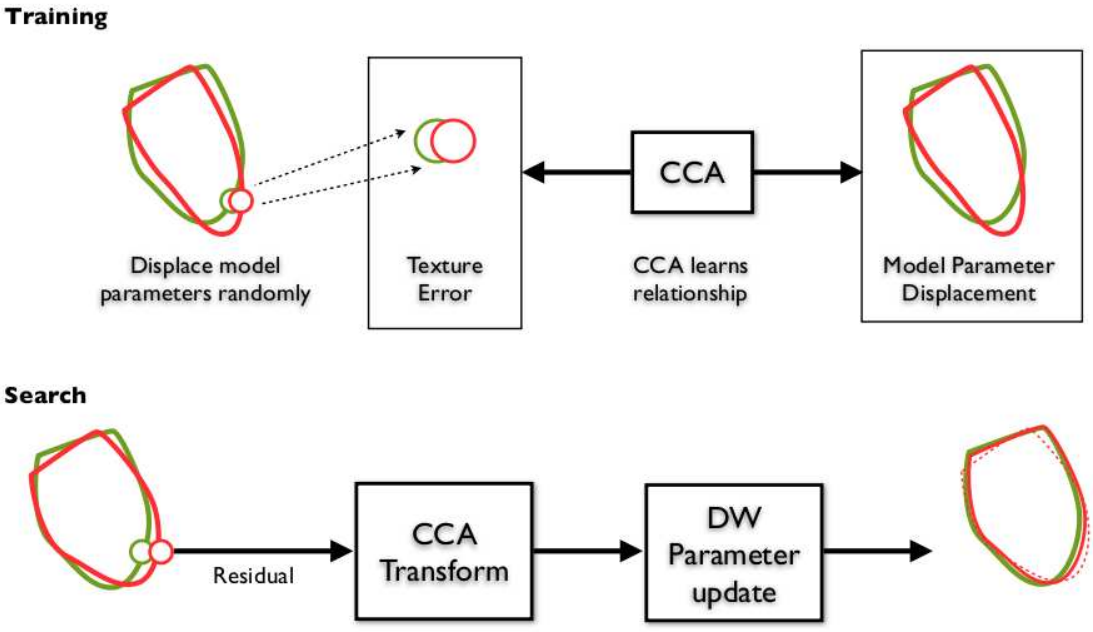


Figure 6.2 — DW-CCA Methodology.

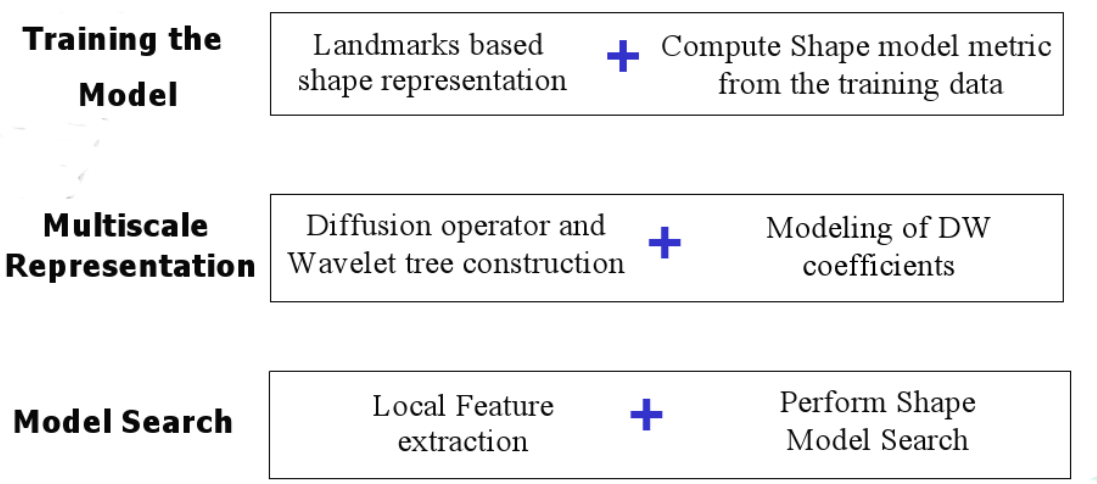


Figure 6.3 — Segmentation Framework.

6.6.1 Image-based Search

To assess the search behavior we compare our method with a standard Gaussian shape model search in an active shape model search approach. We use an even sampling of the object surface, and gradients in the volume as texture description, and a sparse shape model as proposed in [Essafi *et al.*, 2008]. The latter uses a similar appearance model to the one used in this section, and allows for the assessment of the effect of replacing the multivariate Gaussian landmark model, with the diffusion wavelet shape model. The error measure is the mean distance of the model landmarks between standard of reference and segmentation result Figure. 6.4. This gives also an indication of the displacement along the surface, which is relevant if the result is used for navigation. Models are initialized with minimal overlap to the target shape, and the accuracy of the final result was quantified by the mean landmark error between ground truth annotation and search result. For the quantitative comparison, results in Fig.(6.4) clearly show how the diffusion wavelet model outperforms the sparse model with standard parameterization for both anatomical data sets, with for example a mean value of 10.97 voxels for diffusion wavelet model over 13.72 error voxels for the sparse model in the calf data.

The diffusion wavelet model is able to recover the shape with superior accuracy. In the muscle data the standard search approach failed due to the ambiguous texture and local shape variability in large regions of the target shape. In Fig.(6.5) examples for standard, sparse model [Essafi *et al.*, 2008], and multi-scale diffusion wavelet based search are depicted. Furthermore the diffusion wavelet shape prior is significantly better than the local Gaussian prior. Note that one of the important points that distinguishes our methodology from robust ASMs, is that we learn the distribution of both image and shape information during the training phase to optimally exploit the anatomical properties of the data. This is not the case for robust ASMs which for a given sampling consider a subset of the control points according to the observed image during search. In a typical segmentation scenario, our method runs approximately 56 seconds in average with non-optimized code implemented in Matlab 7.5, on a 2GHz DELL Duo Computer with 2Gb RAM.

6.6.2 Local Appearance Features CCA Search

To evaluate the shape representation of the our model, we assess two measures: 1. reconstruction accuracy, and 2. search performance. The goal is to understand how the

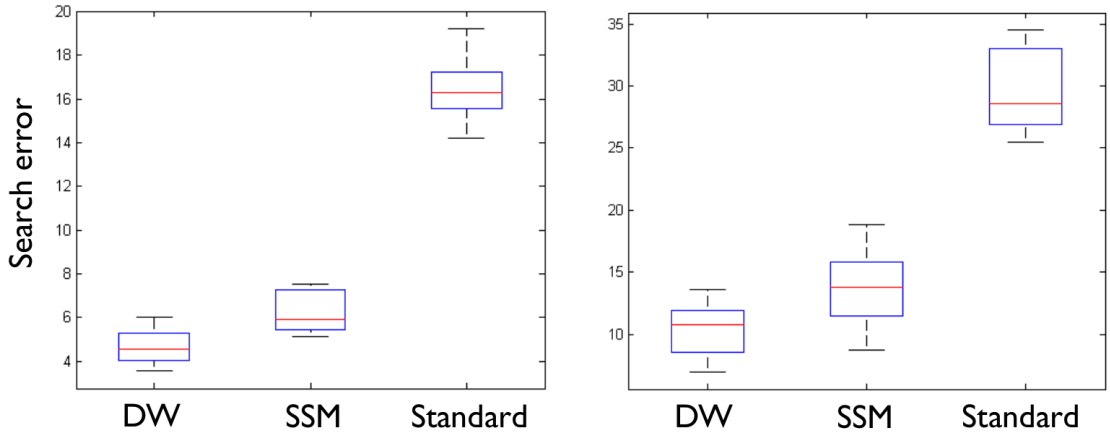


Figure 6.4 — Boxplots of (a) Heart and (b) Calf Search Segmentation. Landmark Reconstruction Error (voxel) after finishing search phase, with comparison between three different search models; (1) our approach, (2) sparse model and (3) standard gaussian model.

hierarchical modeling based on the diffusion kernel affects the reconstruction of the muscle. We compared the reconstruction error for Gaussian shape models, and the proposed diffusion wavelet modeling. The current strategy of combining the diffusion wavelet shape representation and an AFM appearance model strategy outperforms the standard search method, based on an even sampling of the object surface, and gradients in the volumes. The accuracy of the final result was quantified by the mean landmark error between ground truth annotation and search segmentation result of the muscle, as well as the DICE similarity measure coefficient (See Figure. 6.8). The calculated mean distance gives in addition an indication of the displacement along the surface, which is relevant if the result is used for navigation. Models were initialized with minimal overlap to the target shape.

Moreover, one of the main goals is to evaluate how far the diffusion wavelet model is able to detect the local shape variations based on diffusion kernel. In term of quantitative results the reconstruction error between the diffusion wavelet model and the reference model is calculated as an average surface error for all test shapes which gives us 2.1277 voxel for the Gaussian model and 0.1485 for the diffusion wavelet method. In Fig 6.6 the reconstruction of the projected shape model is depicted, this reconstruction starts from the projection of the diffusion wavelet coefficients in respectively the first and the last level, and then extracting a new subset of coefficients from the eigenvectors that constitutes 99% of the variation in the correspondant level.

To keep the variation in reasonable limits, the shape parameters are also restricted to $\pm 3\sqrt{\lambda_i}$. During the reconstruction experiments we were also able to check the effects of each eigenvector of any scale on the rate error.

Regarding the muscle data our approach was able to recover the shape with superior accuracy, in the meanwhile the standard search approach failed due to the ambiguous texture in large regions of the target shape. In Fig 6.7 one can visualize an example of comparison between standard and active feature model search. It is interesting to note that with the help of image support based on local texture descriptors, the method performs better for muscle segmentation due to a restriction to more relevant information being used for regression and fitting.

6.6.3 Appearance Classifier Search

To assess the performance of our approach, we consider a data set that includes 25 CT volumes of the heart, with an approximate voxel spacing of 1.5 mm, for which 90 anatomical standard of reference landmarks, and a set of 1451 control points for the left ventricle was available, in addition to the ground truth segmentation from experts concerning the diastole as well as the systole.

We have run our algorithm in a leave-one-out cross validation fashion. For the diffusion wavelet building part, we obtain 9 diffusion wavelet levels of decomposition for the shape prior. As for the initialization of our framework, we used the mean shape displaced by a random translation of 30 mm.

To evaluate the efficiency of our method, we computed two error measures: (i) the Hausdorff distance revealing the maximum error between the standard of reference and our model reconstruction, as well as (ii) mean distance error of the detected landmarks. In Fig.6.9.a, one can see that the Hausdorff distance error decreases with an increasing number of diffusion wavelet levels used for reconstruction. When we consider the mean reconstruction error over all data, we reach a distance of 2.2313 voxel in the image for the finest level, while as for the coarsest level we obtain 2.7073 voxel. The comparison of detection results for different numbers of levels used during reconstruction can be seen in Fig.6.9. Note that diffusion wavelets have been shown to outperform standard Gaussian models in terms of search error in [Essafi *et al.*, 2009a] on muscle MRI data.

During the search validation experiments, we consider a multi-resolution approach for each landmark patch which goes from 5*5 pixels to 20*20 pixels in 4 steps. We

obtain 200 landmarks candidates, for 15 training hearts and 10 testing examples. Experiments were carried out using Gentle Adaboost, which is adequate to deal with a large number of negative examples as well as the rather limited size of our training set. In the quantitative assessment of the search/segmentation algorithm explained in Sec.6.3, we obtain a lowest error of 4.72 voxel between ground truth and relative segmented volume. In a typical segmentation scenario, the method takes approximately 68 seconds in average through non-optimized code implemented in Matlab 7.5, on a 2GHz DELL Duo Computer with 2Gb RAM. One should note here that we are working toward search in very large data sets, while searching for small complex structures, thus the efficiency of gradient descent of ASM is limited. In an ideal case one would combine the trade off between the reconstruction accuracy and the classification error to choose the best candidate for the search segmentation.

6.7 Contributions

In this chapter we detail three different search schemes for diffusion wavelet shape prior models: 1. A search based on an image based local appearance representation analogous to the search discussed in Chap.4, 2. a search scheme based on a local appearance representation by classifiers, and a hypothesis selection scheme based on the shape prior, and 3. a search scheme based on local features, and the inference of the model parameters based on CCA. In accordance to the results, our proposed strategies have proven to be accurate and suitable for medical images treatment.

Undoubtly, there still is a huge variety of feature sets and classifiers that could be investigated, and in the same aspect other feature extraction and selection techniques could be evaluated. A comparison framework could represent an interesting future direction. The main objective remains to formulate the segmentation problem as an inference task based on data driven hypothesis, and shape prior constraints.

We conclude that the proposed search approaches are able to use the diffusion wavelet shape prior effectively. Together with MDL-based landmark placement during learning, they present a framework to model shape and appearance of anatomical structures. The particular advantages of the proposed approaches are the adaptability to the distribution of informative image content, and a multiscale shape representation, that uses a parameterization learned from the training data to optimally represent the shape variation.

These characteristics are particularly relevant for the computer based analysis of anatomical structures that exhibit complex shape variability, cannot be parameterized sensibly by an a priori chosen reference manifold, and are only partially well discernable in the medical imaging data.

The main concept proposed in this chapter is taking advantages of both the idea of sparse representation and optimal image features for shape representation integrated in a global classification scheme :

Training Based on an initial set of landmarks, we assess first the discriminative power of the appearance at each landmark position (Appearance), and second the reconstruction contribution / shape model redundancy of each landmark.

Search We perform a coarse to fine search - starting with a global search on a small number of highly discriminative landmarks / transitioning to local fine search. Search is classifier based as described above.

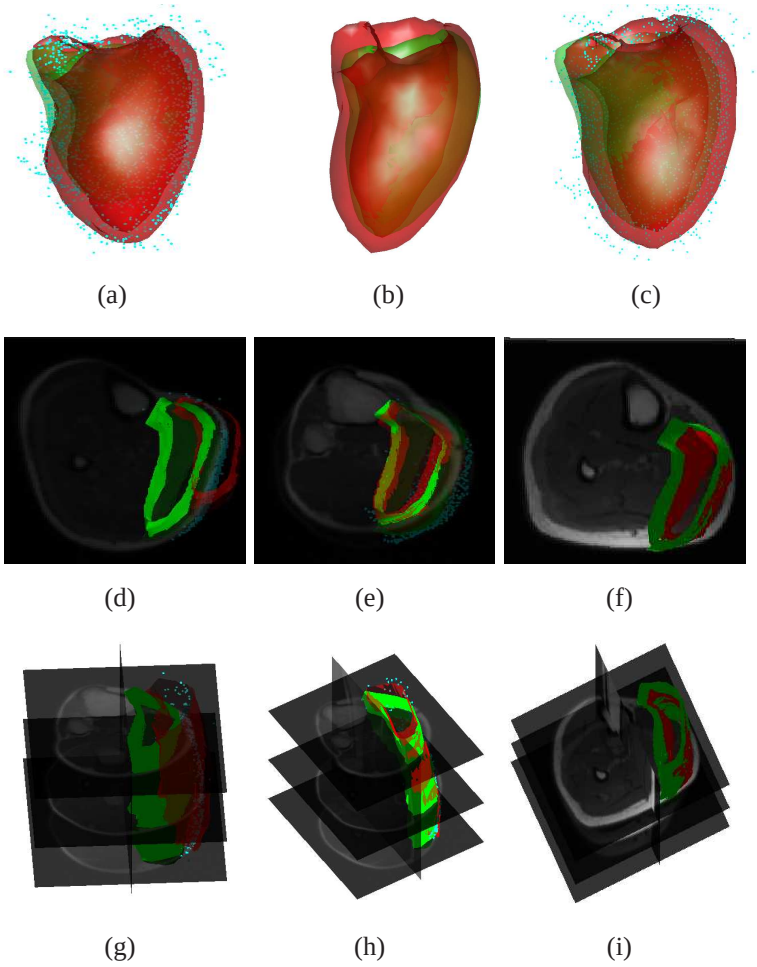
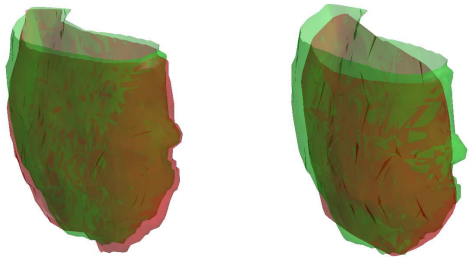
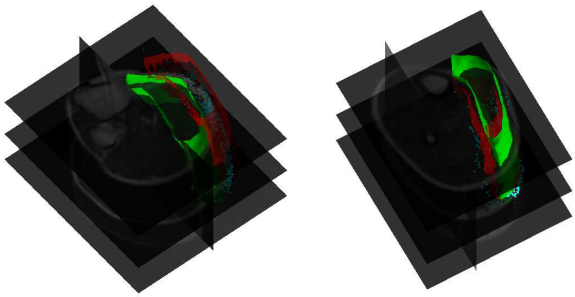


Figure 6.5 — Model search result for Heart muscle (upper row) and Calf muscles (down row). Data in green: standard of reference segmentation, in red: search results. For (a, d, g) standard gradient search approach, while (b, e, h) represent sparse shape models and finally (c, f, i) diffusion wavelet model.



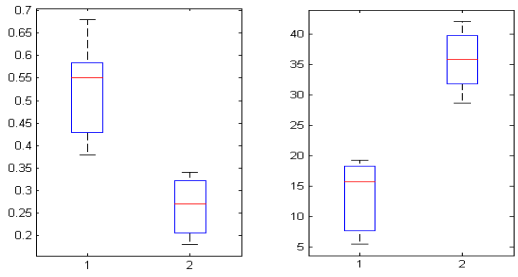
(a) (b)

Figure 6.6 **—** Diffusion Wavelets Model reconstruction. Data, green: standard of reference segmentation, red: reconstruction result for a. finest scale and b. coarsest wavelet scale.



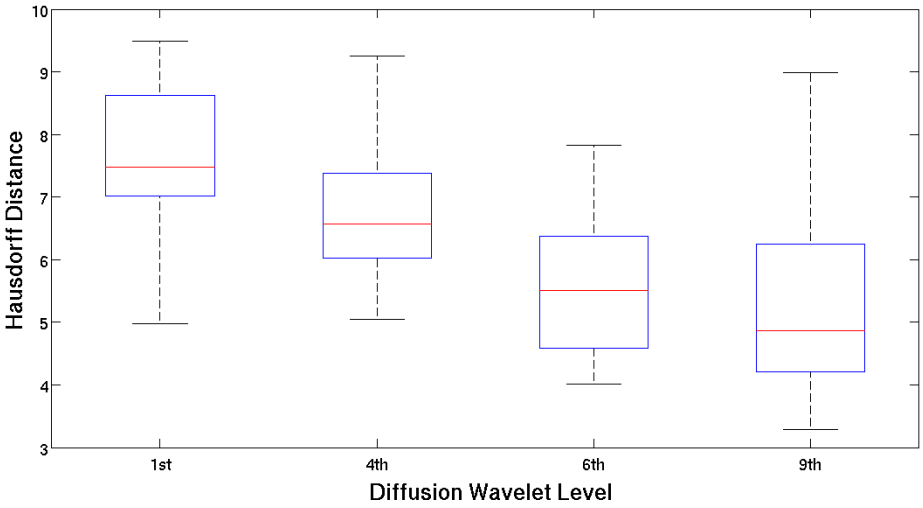
(a) (b)

Figure 6.7 **—** Search Comparison: Model search result for T1 MRI calf , green: "gold standard" segmentation, red: search results. (a) standard gradient search approach and (b) active feature models search.

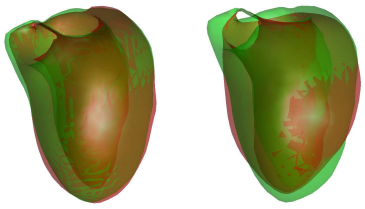


(a) (b)

Figure 6.8 **—** Results of the segmentation using the DW-AFM model. (a) Boxplots of the Dice Similarity Coefficients Measure and (b) landmark error (voxel) after finishing the search phase over the whole data set, with (1) our current approach and (2) standard model.

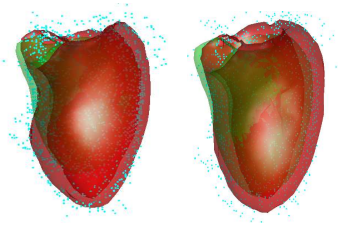


(a)



(b) (c)

Figure 6.9 — Multiscale Diffusion Wavelets Reconstruction. (a) Hausdorff Error Distance (in voxel) of reconstructed heart at each diffusion scale for all data in the training set. (b) Data, green: ground truth segmentation, red: reconstruction result for finest scale and (c) coarsest wavelet scale.



(a) (b)

Figure 6.10 — Model search result for Heart muscle. Ground truth in green, in red: search results. (a) standard Gaussian search approach, and (b) Segmentation based on Image Classifiers.

CHAPTER 7

Robust Sparse Wavelet Enhanced Modeling

“ Le génie, c’est l’enfance retrouvée à volonté “

Baudelaire

7.1 Introduction

In the previous chapters, we have first introduced a model based on sparse shape and appearance representation. Following this, we have explained how one can use diffusion wavelets to represent shape variation in an efficient manner. Both approaches are related to the assumption that one can observe an underlying *topology* of the shape variability during learning. In Chapter. 4, we have used it to find an optimal sub-set of informative landmarks. In Chapter. 5, we have defined the domain on which the wavelets are parameterized according to this topology.

In this chapter we introduce a method for the reconstruction of missing data in a diffusion wavelet shape model framework. By this we bring the two approaches together: the parameterization by means of diffusion wavelets, and the reconstruction of shapes from a subset of landmarks. There are two aspects behind this idea:

1. One can perform robust model fitting, by identifying outliers, and excluding them from reconstruction,

2. One can employ a sparse model based on a sub-sampling of landmarks.

Robustness is a crucial issue in computer vision and shape modeling. It aims to guarantee that shape analysis and reconstruction are robust to noise, occlusions and clutter. We will describe how to detect such outliers during diffusion wavelet shape model search, and how to reconstruct the shape based on the remaining reliable landmarks.

The second novel aspect is the building of Sparse Diffusion Wavelet Models based on the same concept: imputation.

We will name our approach **R-SWAM** standing for a **R**obust **S**parse **WA**velet **M**odel.

The proposed approach aims at offering a high flexibility regarding types of disturbances that can be handled. The robustness and performance of the **R-SWAM** algorithm is evaluated on the same data sets as the previous algorithms, to allow for comparability of the results.

In the first part of this chapter we depict the related work of our method, then in Section. 7.3 we detail the different steps of the **R-SWAM** framework. Finally we validate the method on our medical data set.

7.2 Related Work

An important issue of shape modelling and the capturing of small local deformations typical for muscle data is the modelling of the shape variation. A particular aspect relevant when performing imputation and reconstruction with wavelets, is the lifting scheme. The wavelet lifting scheme consists in decomposing wavelet transforms into a set of stages. A comprehensive introduction is given in the work of [Heaton and Silverman, 2008] who focus on a wavelet lifting scheme based imputation method using sparse representation of a surface in a wavelet/lifting scheme basis. We were indeed inspired by the iterative wavelet imputation described by the authors. The idea of

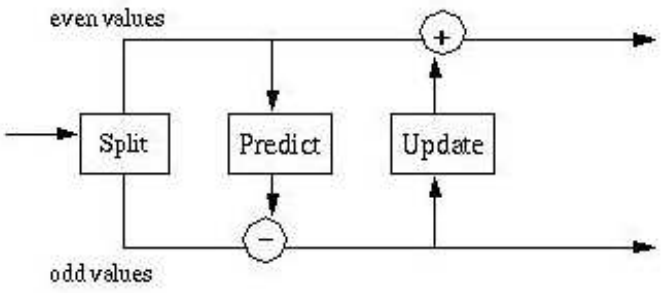


Figure 7.1 **—** Wavelet Lifting Scheme.

The original idea behind the lifting scheme [Schröder and Sweldens, 1995] is to start from one basic and simple multi resolution analysis, and drive the construction of a new and more efficient one. Consequently either the new basis functions are smoother or the wavelets possess more vanishing moments. In fact the lifting enables to build a basis in a fully biorthogonal framework. This scheme guarantees a finite and small support for all the bases, as well as more performance regarding the wavelets.

Several approaches were reported in literature to make Active Appearance Model more robust, such as [Beichel *et al.*, 2005] who developed a robust AAM matching algorithm to resolve problems of gross disturbances in medical data.

7.3 Wavelet Enhanced Sparse Modelling

Our goal is to describe how a sparse model framework can be augmented with diffusion wavelet concept to reduce model complexity in one part, but also to successfully cope with large medical data sets.

7.3.1 Reconstruction of missing landmarks in a diffusion wavelet shape model

Similarly to our previous shape models approaches, we start with a Point Distribution Model, with full landmarks describing our data. Let us define that for m landmarks the positions, $\mathbf{V}_i = \{\mathbf{x}_1^i, \mathbf{x}_2^i, \dots \mathbf{x}_m^i\}$, are known in N training images $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_N$. That is, our shape knowledge comprises $\mathcal{V} = \{\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N\}$, where $\mathbf{x}_j^i \in \mathbb{R}^d$, and we call $\mathbf{V}_i \in \mathbb{R}^{d \times m}$ a shape.

Then, let us assume that we have a subsampled set of the initial landmarks, with;

$$\hat{\mathcal{V}} = \{\hat{\mathbf{V}}_1, \hat{\mathbf{V}}_2, \dots, \hat{\mathbf{V}}_n\}. \tag{7.1}$$

where $\hat{\mathbf{V}}_i \in \mathbb{R}^{d \times m'}$ are the representations of the full shapes V_i in the training set, with $m' \ll m$. $\hat{\mathbf{V}}_i$ consists of a sub set of the landmarks defining the shape. This can be the result of one of the followings causes:

1. A sub-sampling procedure as described in Section. 4.5.

2. Missing landmarks due to corrupted, noisy or outliers data.

The main goal now is to try to predict $\mathbf{V}_i$ from $\hat{\mathbf{V}}_i$, however this time through the use of the diffusion wavelet models.

As we have already built the the diffusion wavelet tree Φ of our complete training set (as described in Section. 5.4), we use it to represent the subsampled test volume. We calculate the diffusion wavelet coefficient $\Gamma_{\hat{\mathbf{S}}_\mathbf{i}}$ on the deviation $\hat{\mathbf{S}}_\mathbf{i}$ from the mean of the aligned shapes, Let us denote by F_0 the set of fixed landmarks and M_0 the missing ones in the subsampled shape, corresponding to;

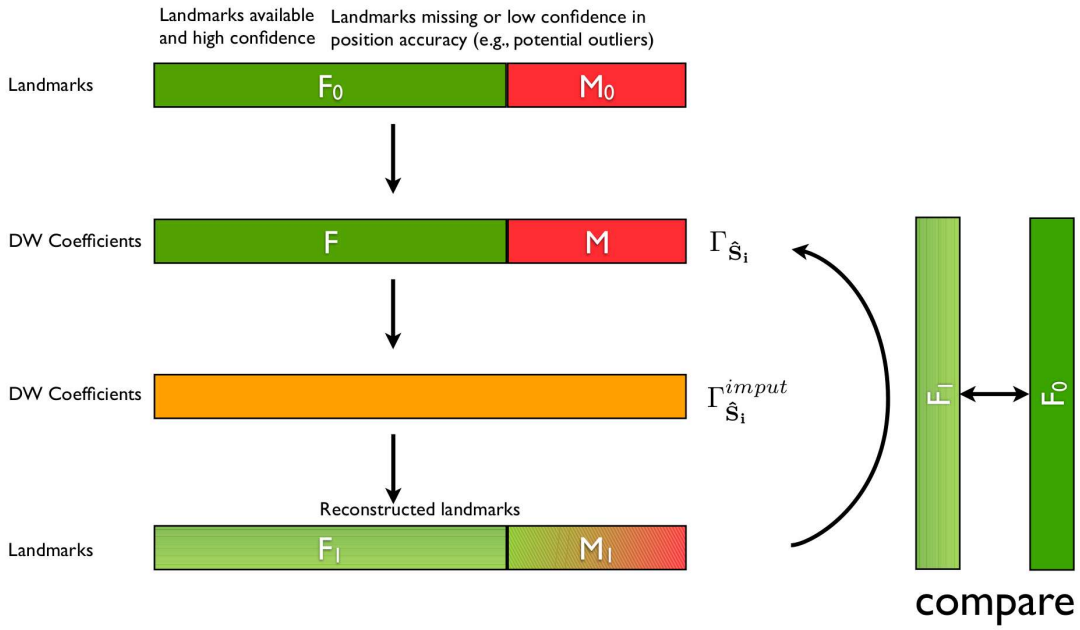


Figure 7.2 — Overview of the **R-SWAM** Framework.

$$\hat{\mathbf{S}}_i = F_0 \cup M_0 \tag{7.2}$$

and then obtain the following diffusion wavelet coefficients for the example $\hat{\mathbf{S}}_\mathbf{i}$,

$$\Gamma_{\hat{\mathbf{S}}_\mathbf{i}} = \Phi^{-1}\hat{\mathbf{S}}_\mathbf{i} \tag{7.3}$$

Afterward, we imputate the sparse diffusion wavelet coefficients, $\Gamma^{imput}_{\hat{\mathbf{S}}_\mathbf{i}}$, by projecting them into the original PCA space of the full diffusion wavelets training, i.e., the eigenvectors $\Sigma = \{\sigma_j\}_{j=1\ldots K}$, and the corresponding eigenvalues $\Lambda = \{\lambda_j\}_{j=1\ldots K}$.

Once we finish the imputation step, we are now able to reconstruct the subsampled shape as following:

$$\hat{\mathbf{S}}^{\mathbf{rec}}_\mathbf{i} = \mathbf{\Phi}\mathbf{\Gamma}^{\mathbf{imput}}_{\hat{\mathbf{S}}_\mathbf{i}} \tag{7.4}$$

However to optimize the reconstruction, we will take into consideration the new reconstructed values F_1 relative to the positions of the original fixed landmarks F_0.

$$\hat{\mathbf{S}}_{\mathbf{i}}^{\mathbf{rec}} = \mathbf{F_1} \cup \mathbf{M_1} \tag{7.5}$$

We will iterate then this process, until the reconstruction error between the original fixed landmarks F_0 and the fixed reconstructed ones F_1 reaches a certain threshold ϵ.

Finally the reconstructed shape is obtained by;

$$\hat{V}_i^p = \bar{V}^p + \Phi\Gamma_{\hat{\mathbf{S}}_{\mathbf{i}}}^{imput} \tag{7.6}$$

7.3.2 Sparse Diffusion Wavelet Framework

Instead of local multivariate Gaussians, as in Chapter. 4, the shape variation observed in a training set of shapes is represented with a diffusion wavelet shape models.

The approach proceeds in the following steps:

1. Consider a set of training volumes, and correspondences across the examples for a set of landmarks.
2. Learn a diffusion wavelet model from the complete set of landmarks
3. Choose a sparse sub set of landmarks modeling analogously to the approach described in Chapter. 4.
4. Estimate diffusion wavelet coefficients based on the sparse point set of during search volume.
5. Perform the robust reconstruction of the entire model as described in Section.7.3.1.

The above described framework is equivalent to determine the diffusion wavelet coefficients from an incomplete set of points. However those points will not be chosen randomly, but established through the Sparse Shapes Model. Obviously a robust version could always be considered, starting from RANSAC points (”RANdom SAmple Consensus"), one investigates how well generated diffusion wavelets fits the ground truth and then we can choose those which have the most support i.e. which estimate most points.

7.4 Experimental Results

In this experimental validation section, we evaluate the multiscale shape prior based on a shape reconstruction task form a subset of landmarks. The prior is learned during the training phase from a set of examples, and the full set of landmarks. Afterward we fit an incomplete test shape with the prior so as to estimate how close a reconstructed test shape is to the ground truth.

We then validate our approach on a set of 25 CT volumes of the heart, with an approximate voxel spacing of 1.5 mm, for which 90 anatomical standard of reference landmarks, and a set of 726 control points for the left ventricle was available, also with available ground truth segmentation from experts concerning the diastole and the systole.

We test the diffusion wavelet model on the sparse set of landmarks obtained as described in Chapter. 4. We calculate the residual error between full diffusion wavelet coefficients and altered ones, the result is shown in Figure. 7.3.

We compute the mean error distance in order to calculate the discrepancy between the ground truth segmentation and the reconstructed result. This distance will then measure boundaries of both volumes. In Table.(7.1), we report errors of diffusion wavelet models and R-SWAM technique.

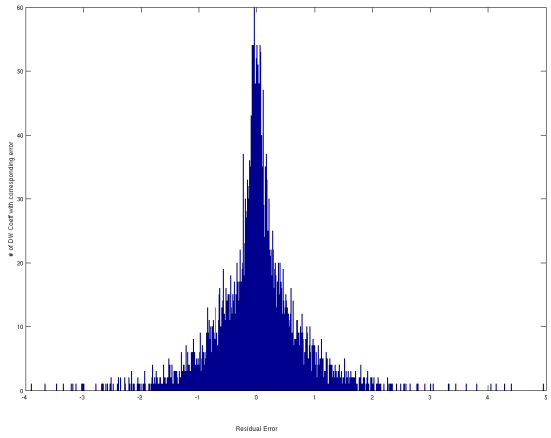


Figure 7.3 — Selection of residuals and their effect on diffusion wavelet projection

In Figure. 7.4 different iterations of the sparse diffusion wavelet reconstruction of a subsampled heart volume is shown. One can notice that the predicted shape of the

	R-SWAM	Diffusion Wavelet
10% Missing Landmarks	0.264	6.244
50% Missing Landmarks	0.313	9.328

***Table 7.1* —** Comparative table results between Diffusion Wavelet and R-SWAM, over the reconstruction landmark error between the ground truth and the reconstructed sparse volume of the left ventricle muscle.

R-SWAM is improving with every iteration, leading to the minimization of the reconstruction error rate. It's interesting to note that regarding the figure results as well as numeric values, that our robust technique is able to retrieve a sparse landmarks shape, in an even better way than conventional methods.

To investigate the influence of different amounts of missing data on the robust procedure, we performed a comparison of R-SWAM with the standard DW model on different rate of sparse missing landmarks which were replaced by noise. Clearly the standard DW model fails in reconstructing the shape. The result can be shown in Figure. 7.5. As demonstrated in the state of the art, experiments prove that standard Gaussian model does not handle missing landmarks or outliers well. Obtained results demonstrate that robust diffusion can overcome this kind of drawbacks.

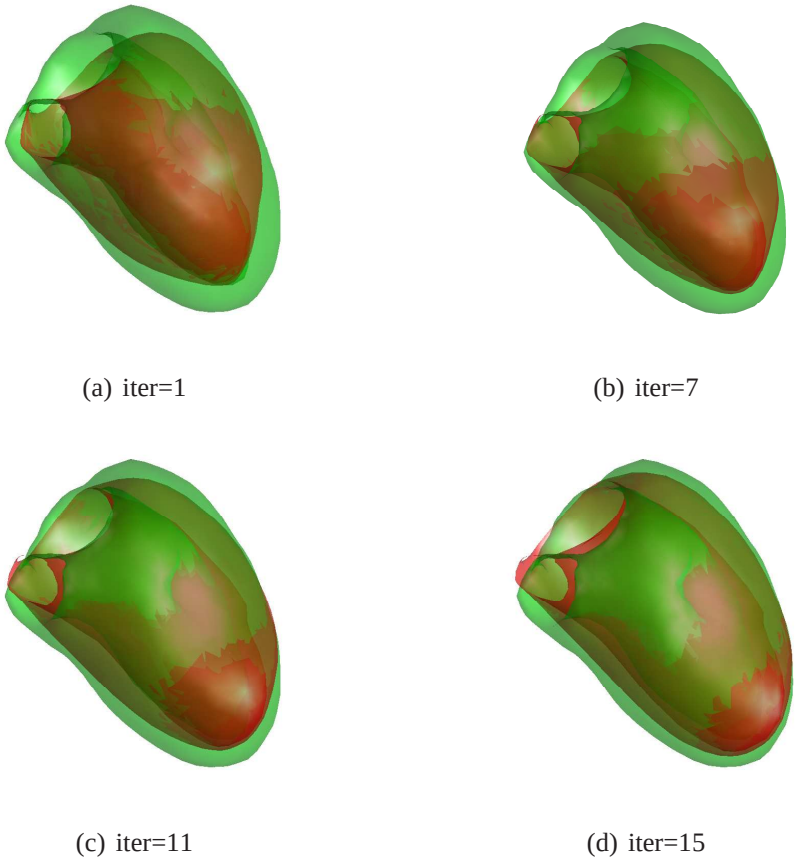
Similarly to the previous experimental validation, this algorithm runs on a Matlab 7.5, on a 2GHz DELL Duo Computer with 2Gb RAM.

7.5 Contributions

In this chapter, we proposed to extend the framework proposed in the two previous chapters. We introduce a reconstruction algorithm, that can estimate the diffusion wavelet shape model coefficients from a subset of landmarks. That is the model can deal with missing landmarks, that are either due to detected outliers, or to a sparse subsampling of the landmarks analogously to Chapter. 4. By using both the diffusion wavelet and sparse modeling as shape representation, we are able to take advantage of the decomposition space on the one hand and sparse decomposition on the other hand.

A part from handling the missing landmarks issue, and offering hierarchical decomposition, our technique is general in the sense that it offers the advantage of being suit-

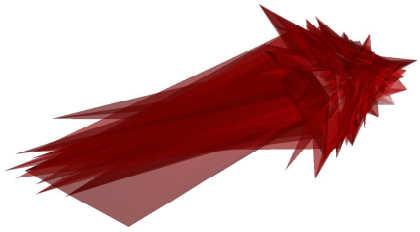
able to any medical acquisition modality technique and to any anatomical structure.



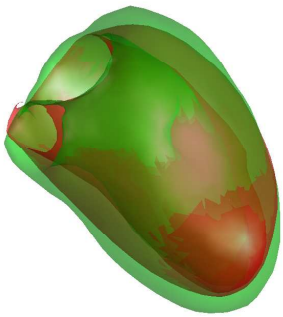
(a) iter=1 (b) iter=7

(c) iter=11 (d) iter=15

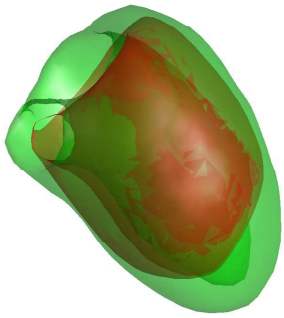
Figure 7.4 — Several Iterations of Sparse Diffusion Wavelet Reconstruction.



(a) Non robust Diffusion Wavelet Model



(b) 10% of Missing Landmarks



(c) 50% of Missing Landmarks

Figure 7.5 — (a) Non robust Diffusion Wavelet Model Reconstruction on noisy data. Comparison between R-SWAM reconstruction for respectively (b) 10% and (c) 50% of sparse missing landmarks.

Conclusion

“If we knew what it was we were doing, it would not be called research, would it?”

Albert Einstein

7.6 Contributions

In this dissertation, two knowledge-based segmentation frameworks for medical imaging data are proposed. The frameworks tackle questions regarding the optimal modelling of anatomical structures and an efficient search in ambiguous image data: 1. They cope with heterogeneous distributions of salient image features that can be used during localization and segmentation. 2. They use the shape variability and its redundancies observed in the training set optimally to build compact and sparse models, 3. They learn an adaptive parameterization of the shape variation modeling domain, that can is learned form the training data, instead of being imposed following a priori assumptions. The medical background is the segmentation of muscles in MRI data, a problem that poses the above mentioned challenges. The two approaches are:

- **Sparse shape models** adapt to heterogeneous distributions of redundancy in the shape variation and sparsely distributed distinctive texture in the data. In contrast to existing approaches they use the statistical modeling and texture behavior to derive a sparse representation and reconstruction mechanism.

- **Diffusion wavelet shape priors** are a segmentation framework based on diffusion wavelets and local appearance classifiers. The conjunction of the diffusion wavelet constraint with a search method based on a GentleBoost classifier leads to an effective segmentation scheme.

Sparse shape models obtain a sparse model of objects or anatomical structures that takes the local statistical modelling and appearance behavior in the training set into account. The model based segmentation is based on a sparse set of landmarks that can be uniquely identified in new data during search. The model is built based on a set of training examples for which expert annotations are available. It integrates knowledge about local appearance, shape variability, and the ambiguity of image data to achieve a segmentation performance equal to or superior to a medical expert.

During search for structures in new data, the model is used with a patch based local appearance representation to locate and segment objects. The proposed representation can be employed with other shape modeling and search methods. The formulation based on model compactness makes a transfer of the shape behavior mapping to other models straightforward. A comparison of the reconstruction and search behavior with standard

shape models that neglect these properties and represent objects regardless of the reliability and complexity of shape and texture behaviour in the training population shows clear advantages of this strategy.

Diffusion wavelet shape priors are a multi-scale shape modeling framework based on a diffusion wavelets shape representation. The method takes advantage of the subtle inter-dependencies in training data, by clustering coefficients based on correlation, and representing the topology of the structure by a diffusion kernel, instead of a fixed predefined manifold. We are using the orthomax criterion which is suitable for building sparse representations - particularly relevant in the case of the regions and pathologies studied in the thesis - leading to localized modes of variation an optimal subdivision of the shape parameterization. A segmentation framework based on diffusion wavelets and local appearance classifiers allows to use the model during search for anatomical structures in new data. The conjunction of the diffusion wavelet constraint with a search method based on a GentleBoost classifier leads to an effective segmentation scheme that can deal with ambiguous appearance and complex structures. We have shown that in the context of anatomical structures, the diffusion wavelet transformation is able to accurately and efficiently detect the locations and spatial scales of shape variations.

Finally a scheme for the reconstruction of partial shapes by means of the diffusion wavelet shape prior is introduced. It is the basis for robust search approaches, and for a sparse subsampling of the most informative landmarks and a subsequent reconstruction of the entire shape. The proposed approaches contribute a more flexible shape and appearance model learning framework to the line of research largely initiated by the introduction of active shape models. They make use of the information learned from the training set, by learning not only the variation of the observed shapes, and appearance, but also an optimal parameterization of the shape model. Diffusion wavelet shape priors can be viewed as an extension to the work on spherical wavelet shape models [Nain *et al.*, 2007] towards shapes without a topology known a priori. Instead of defining a topology by a reference manifold, we learn a diffusion kernel that defines locality. The validation of the localization of two complex medical data sets shows promising results indicating the advantage of using such a learned model parameterization.

7.7 General Limitations

The elements presented in this thesis carry on certain strengths and consistency, but still have some limitations. One critical point in our framework consists in registering the training volumes as a pre-processing step. We still have to investigate the sensitivity of our approach to initialization variability, and capture range in large images, and its performance when segmenting large compounds of anatomical structures. These questions are more difficult in the case of the calf muscle data, where more investigation is needed towards extending to the entire set of muscles..

7.8 Future Research

There remain challenges to the approaches presented in this thesis. For future work one can focus on the determination of the sparse model complexity, that takes the limited training data into account to estimate a feasible number of parameters, and an integration with model learning approaches, that learn the locations in non-annotated data and a weakly- or unsupervised manner. Such experiments would require more data to achieve and validate further improvement.

More efficient optimization techniques could be used in conjunction with our priors to obtain a flexible and powerful paradigm representing/inferring shapes of arbitrary topologies. The coarse to fine search, through objective function optimization could also be seen more in depth. And still in the part of diffusion wavelets one can think about different non linear prior estimation instead of pca/orthomax, such as kernel PCA for example as described in [Dambreville *et al.*, 2008]. Another attractive perspective is the combination of the Diffusion Wavelet with the Discrete optimization presented in [Besbes *et al.*, 2009], where the authors explore Discrete MRF (Markov Random Fields) for knowledge based segmentation.

A major line of future work is the integration of the proposed methods, in a pathology assessment framework for myopathies, for which the entire group of calf muscles has to be modeled and assessed. We presented the application of the orthomax principle on our data. Currently the technique presented is limited to the varimax criterion. We assumed that similarly to [Leung and Bosch, 2007b] the results under other criteria will be analogous, because the reparametrizations are restricted by the variations of the training set.

However in the future one would like to compare different orthomax criteria, and assess their influence on the shape modelling.

An interesting option is to integrate segmentation and fiber classification based on DTI study [Neji *et al.*, 2009] to a unified framework that allows for the computer-aided assessment of the disease status and the comparison across subjects during follow up examinations. The latter work refers to the application of the unsupervised learning methods developed for muscle fiber clustering and feature extraction on diseased muscle data. These methods were used to find characteristic features in the fiber trajectography data, with high differentiation power between healthy and diseased tissue.

The proposed research program by AFM (French Association against Myopathy), along with the achieved results, hold a great promise for the future. The project has great potential value for the purpose of helping the diagnosis of muscular diseases, evaluating in vivo the impact of treatments of myopathies by a non invasive method, or even more generally exploring muscle structure through a novel, and most importantly non invasive perspective. The prime aim remains to achieve the best possible accuracy for clinical application, the quantitative assessment and treatment support of myopathies.

Publications

- **Articles in International Conference Proceedings**

  - C. Wang, O. Teboul, F. Michel, S. Essafi and N. Paragios. "3D Knowledge-based Segmentation using Pose-Invariant Higher-Order Graphs". In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Beijing (China), September 2010. Springer.

  - S. Essafi, G. Langs and N. Paragios. "Hierarchical 3D Diffusion Wavelet Shape Priors". In 12th IEEE Proceedings of International Conference on Computer Vision (ICCV), Kyoto (Japan), October 2009. [Essafi *et al.*, 2009b]

  - S. Essafi, G. Langs and N. Paragios. "Left Ventricle Segmentation Using Diffusion Wavelets and Boosting". In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pages 919-926. London (UK), September 2009. Springer. [Essafi *et al.*, 2009c]

  - S. Essafi, G. Langs, J-F. Deux, A. Rahmouni, G. Bassez and N. Paragios. "Wavelet-Driven Knowledge-Based MRI Calf Muscle Segmentation". In Proceedings of the International Symposium on Biomedical Imaging (ISBI) : From Nano to Macro. Boston (MI, USA), June 2009. IEEE. [Essafi *et al.*, 2009a] **Oral Presentation**

– S. Essafi, G. Langs and N. Paragios. “Sparsity, Redundancy, Optimal Image Support: an axiomatic approach to knowledge-based segmentation”. In Proceedings of Computer Vision and Pattern Recognition (CVPR). Alaska (USA). June 2008. IEEE. [Essafi *et al.*, 2008]

- **Abstracts**

  – R. Neji, S. Essafi, G. Langs, G. Fleury, A. Sotiras, J-F. Deux, A. Vignaud, A. Rahmouni, G. Bassez, P. Carlier and N. Paragios.“Perspectives in Myopathy Assessment: Virtual Biopsy and the Potential of Medical Image Analysis using Emerging Modalities“. In 6th International Workshop for Musculoskeletal and Neural Interactions, Cologne (Germany). May 2008.

  – S. Essafi, G. Langs, G. Bassez, J-F. Deux, A. Vignaud, A. Rahmouni and N. Paragios. ”Sparse Shape Models with Optimal Image Support for MRI Calf Muscle Segmentation“. In 3rd International Congress of Myology, Marseille (France). May 2008.

- **Technical Report**

  – S. Essafi, G. Langs, J-F. Deux, A. Rahmouni, G. Bassez and N. Paragios. ”Local Appearance Knowledge and Shape Variation Models for Muscle Segmentation“. Galen Group - INRIA Saclay Ile de France. May 2009.

  – S. Essafi, G. Langs and N. Paragios. ”Sparse Notes over muscle segmentation“. Galen Group - INRIA Saclay Ile de France. May 2008.

Bibliography

[Abi-Nahed *et al.*, 2006] J. Abi-Nahed, M.P. Jolly, and G. Yang. Robust active shape models: A robust, generic and simple automatic segmentation tool. *Lecture Notes in Computer Science*, 4191:1, 2006.

[Antoine and Vandergheynst, 2009] J. P. Antoine and D. Rosca P. Vandergheynst. Wavelet transform on manifolds: old and new approaches. *Applied and Computational Harmonic Analysis*, 2009.

[Basser and Pierpaoli, 1998] P. J. Basser and C. Pierpaoli. A simplified method to measure the diffusion tensor from seven mr images. *Magn Reson Med*, 39:928–934, 1998.

[Bassett, 2002] D.R.Jr Bassett. Scientific contributions of AV Hill: exercise physiology pioneer. *Journal of applied physiology*, 93(5):1567, 2002.

[Beichel *et al.*, 2005] R. Beichel, H. Bischof, F. Leberl, and M. Sonka. Robust active appearance models and their application to medical image analysis. *IEEE Transactions on Medical Imaging*, 24(9):1151–1169, 2005.

[Belkin and Niyogi, 2003] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation. MIT Press*, 15:1373–1396, 2003.

[Besbes *et al.*, 2009] A. Besbes, N. Komodakis, G. Langs, and N. Paragios. Shape priors and discrete mrfs for knowledge-based segmentation. In *CVPR*, pages 1295–1302, 2009.

[Bihan *et al.*, 2001] D. L. Bihan, J.F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of magnetic resonance imaging*, 13:534–546, 2001.

[Bizais *et al.*, 1995] Y. Bizais, C. Barillot, and R. Di Paola. *Information Processing in Medical Imaging*. Computational Imaging and Vision. Kluwer Academic Publishers, 1995.

[Blemker *et al.*, 2007] S. S. Blemker, D. S. Asakawa, G. E. Gold, and S. L. Delp. Image-based musculoskeletal modeling: applications, advances, and future opportunities. *Journal of Magnetic Resonance Imaging*, 25(2):441–451, 2007.

[Bookstein, 1989] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *TPAMI: IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[Bookstein, 1997a] F. L. Bookstein. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis*, 1(3):225–243, 1997.

[Bookstein, 1997b] F. L. Bookstein. *Morphometric tools for landmark data: geometry and biology*. Cambridge Univeresity Press, 1997.

[Bookstein, 1997c] F.L. Bookstein. Two shape metrics for biomedical outline data: Bending energy, procrustes distance, and the biometrical modeling of shape phenomena. In *Proceedings International Conference on Shape Modeling and Applications*, pages 110 –120, 1997.

[Boykov and Jolly, 2001] Y. Boykov and M.P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in ND images. In *International Conference on Computer Vision*, volume 1, pages 105–112, 2001.

[Brammer, 1998] M. J. Brammer. Multidimensional wavelet analysis of functional magnetic resonance images. *Human Brain Mapping*, 6:378–382, 1998.

[Brechbuehler *et al.*, 1995] C. Brechbuehler, G. Gerig, and O. Kubler. Parametrization of closed surfaces for 3-D shape description. *Computer Vision and Image Understanding*, 61(2):154–170, 1995.

[Butz *et al.*, 2003] T. Butz, P. Hagmann, E. Tardif, R. Meuli, and J. Thiran. A new brain segmentation framework. In *Proceedings of MICCAI'03: International Conference on Medical Image Computing and Computer Assisted Intervention*, volume 2879, pages 586–593, 2003.

[Charpiat *et al.*, 2005] G. Charpiat, O. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, 2005.

[Chui and Rangarajan, 2000] H. Chui and A. Rangarajan. A new algorithm for non-rigid point matching. In *Proceedings CVPR'00: IEEE Computer Society Conference on Computer Visionand Pattern Recognition*, pages 2044–2051, 2000.

[Chung, 1997] F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

[Cohen and Cohen, 1993] L.D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1131–1147, 1993.

[Cohen, 1991] L. D. Cohen. On active contour models and ballons. *CVGIP: Image Understanding*, 53:211–218, 1991.

[Coifman and Lafon, 2006] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21:5–30, 2006.

[Coifman and Maggioni, 2006] R. R. Coifman and M. Maggioni. Diffusion wavelets. *Applied and Computational Harmonic Analysis*, 21:53–94, 2006.

[Coifman *et al.*, 2005] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, F. Warner, and S. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. In *PNAS: Proceedings of the National Academy of Sciences*, pages 7426–7431, 2005.

[Cootes and Kittipanya-ngam, 2002] T. F. Cootes and P. Kittipanya-ngam. Comparing variations on the active appearance model algorithm. In *Proceedings of BMVC'02*, volume 2, pages 837–846, 2002.

[Cootes and Taylor, 1995] T. F. Cootes and C.J. Taylor. Combining point distribution models with shape models based on finite-element analysis. *Image and Vision Computing*, 13(5):403–409, June 1995.

[Cootes and Taylor, 2001a] T. F. Cootes and C. J. Taylor. Statistical models of appearance for computer vision. *World Wide Web Publication, February*, 2001.

[Cootes and Taylor, 2001b] T. F. Cootes and Chris Taylor. Constrained active appearance models. In *Proceedings International Conference on Computer Vision ICCV*, volume 1, pages 748–754, 2001.

[Cootes *et al.*, 1992] T. F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Training models of shape from sets of examples. In *Proceedings of BMVC'92: British Machine Vison Conference*, pages 266–275, 1992.

[Cootes *et al.*, 1994a] T. F. Cootes, A. Hill, C.J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12(6):355–366, 1994.

[Cootes *et al.*, 1994b] T. F. Cootes, C. J. Taylor, and A. Lanitis. Active shape models: Evaluation of a multi-resolution method for improving image search. In *Proceedings of BMVC'94: British Machine Vision Conference*, volume 336. Citeseer, 1994.

[Cootes *et al.*, 1995] T. F. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61:38–59, 1995.

[Cootes *et al.*, 1998a] T. F. Cootes, G. Edwards, and C.J. Taylor. A comperative evaluation of active appearance model algorithms. In *Proceedings of BMVC'98: British Machine Vision Conference*, volume 2, pages 680–689, 1998.

[Cootes *et al.*, 1998b] T. F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. In *Proceedings of ECCV'98: European Conference on Computer Vision*, pages 484–498, 1998.

[Cootes *et al.*, 2001] T. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:681–685, 2001.

[Cootes *et al.*, 2005] T. F. Cootes, C.J. Twining, V. Petrović, and C.J. Taylor. Groupwise construction of appearance models using piece-wise affine deformations. In *Proc. of BMVC'05*, 2005.

[Cossu and Clemens, 2001] G. Cossu and P.R. Clemens. Gene and cell therapy for primary myopathies. *The Muscular Dystrophies*, page 261, 2001.

[Cremers and Rousson, 2007] D. Cremers and M. Rousson. *Deformable Models.Theory and Biomaterial Applications*, chapter Efficient Kernel Density Estimation Of Shape And Intensity Priors For Level Set Segmentation, pages 447–460. Jasjit S. Suri and Aly A. Farag, 2007.

[Cremers *et al.*, 2002] D. Cremers, T. Kohlberger, and C. Schnorr. Nonlinear shape statistics in mumford-shah based segmentation. In *Proceedings of ECCV'02: European Conference on Computer Vision*, volume 2, pages 93–108, 2002.

[Criminisi *et al.*, 2009] A. Criminisi, J. Shotton, and S. Bucciarelli. Decision forests with long-range spatial context for organ localization in ct volumes. Technical report, Microsoft Research, Cambridge, UK, 2009.

[Dambreville *et al.*, 2008] S. Dambreville, Y. Rathi, and A. R. Tannenbaum. A Framework for Image Segmentation Using Shape Models and Kernel Space Shape Priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(8):1385–1399, 2008.

[Daniel *et al.*, 1976] J. W. Daniel, W. B. Gragg, L. Kaufman, and G. W. Stewart. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Mathematics of Computation*, 30(136):772–795, 1976.

[Davatzikos *et al.*, 2003] C. Davatzikos, X. Tao, and S. Dinggang. Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, 22:414–423, 2003.

[Davies *et al.*, 2002a] R. H. Davies, C. Twining, T. F. Cootes, and C. J. Taylor. A minimum description length approach to statistical shape modelling. *IEEE Transactions on Medical Imaging*, 21:525–537, 2002.

[Davies *et al.*, 2002b] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. 3D statistical shape models using direct optimization of description length. In *Proceedings of ECCV'02: European Conference on Computer Vision*, pages 3–20, 2002.

[Davies, 2002] R. H. Davies. *Learning Shape: Optimal Models for Analysing Natural Variability.* PhD thesis, Division of Imaging Science and Biological Engineering, University of Manchester, Manchester, England, 2002.

[De Bruijne and Nielsen, 2004] M. De Bruijne and M. Nielsen. Shape particle filtering for image segmentation. *Lecture Notes in Computer Science*, pages 168–175, 2004.

[Delp and Loan, 1995] S.L. Delp and J.P. Loan. A graphics-based software system to develop and analyze models of musculoskeletal structures. *Computers in Biology and Medicine*, 25(1):21–34, 1995.

[Deschamps and Cohen, 2001] T. Deschamps and L. Cohen. Fast extraction of minimal paths in 3d images and application to virtual endoscopy. In *Medical Image Analysis*, volume 5, pages 281–299, December 2001.

[Donner *et al.*, 2006] R. Donner, M. Reiter, G. Langs, P. Peloschek, and Horst Bischof. Fast active appearance model search using canonical correlation analysis. *TPAMI: IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10):1690 – 1694, October 2006.

[Donoho, 2006] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[Dornaika and Ahlberg, 2004] F. Dornaika and J. Ahlberg. Fast and reliable active appearance model search for 3-d face tracking. *IEEE Trans. on Systems, Man and Cybernetics - Part B Cybernetics*, 34(4):1838–1853, August 2004.

[Duncan and Ayache, 2000] J.S. Duncan and N. Ayache. Medical image analysis: progress over two decades and the challenges ahead. *TPAMI: IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:85–106, 2000.

[Duta and Sonka, 1997] N. Duta and M. Sonka. Segmentation and interpretation of MR brain images using an improved knowledge-based active shape model. *Lecture Notes in Computer Science*, pages 375–380, 1997.

[Edwards *et al.*, 1998] G. Edwards, C.J. Taylor, and T.F. Cootes. Interpreting face images using active appearance models. In *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, pages 300–305, 1998.

[Emery, 1993] A.E.H. Emery. *Duchenne muscular dystrophy*. Oxford, 1993.

[Ericsson, 2006] A. Ericsson. *Automatic Shape Modelling with Applications in Medical Imaging*. PhD thesis, Mathematics LTH, Lund University, sep 2006.

[Essafi *et al.*, 2008] S. Essafi, G. Langs, and N. Paragios. Sparsity, redundancy and optimal image support towards knowledge-based segmentation. In *Proceedings CVPR'08: IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[Essafi *et al.*, 2009a] S. Essafi, G. Langs, J-F. Deux, A. Rahmouni, G. Bassez, and N. Paragios. Wavelet-driven knowledge-based mri calf segmentation. In *Proceedings of ISBI'09: IEEE International Symposium on Biomedical Imaging*, 2009.

[Essafi *et al.*, 2009b] S. Essafi, G. Langs, and N. Paragios. Hierarchical 3d diffusion wavelets shape priors. In *Proceedings of ICCV'09 :IEEE International Conference in Computer Vision*, 2009.

[Essafi *et al.*, 2009c] S. Essafi, G. Langs, and N. Paragios. Left ventricle segmentation using diffusion wavelets and boosting. In *Proceedings of MICCAI'09: International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 919–926, 2009.

[Fernandez and Hunter, 2005] J. W. Fernandez and P. J. Hunter. An anatomically based patient-specific finite element model of patella articulation: towards a diagnostic tool. *Biomech Model Mechanobiol*, 4:20–38, 2005.

[Filler *et al.*, 2004] A. G. Filler, K. R. Maravilla, and J. S. Tsuruda. MR neurography and muscle MR imaging for image diagnosis of disorders affecting the peripheral nerves and musculature. *Neurologic Clinics*, 22(3):643–682, 2004.

[Florin *et al.*, 2007] C. Florin, N. Paragios, G. Funka-Lea, and J. Williams. Liver segmentation using sparse 3d prior models with optimal data support. *Proceedings of IPMI'07: Information Processing in Medical Imaging*, pages 38–49, 2007.

[Frangi *et al.*, 2001] A. F. Frangi, W. J. Niessen, and M. A. Viergever. Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE Transactions on Medical Imaging*, 20(1):2–5, 2001.

[Freund and Schapire, 1996] Y. Freund and R. E. Schapire. Experiments with a new boosting algorithm. In *Machine Learning International Conference*, pages 148–156. Citeseer, 1996.

[Friedman *et al.*, 2000] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: A statistical view of boosting. *Annals of statistics*, pages 337–374, 2000.

[Fripp *et al.*, 2007] J. Fripp, S. Crozier, S.K. Warfield, and S. Ourselin. Automatic extraction of the bone cartilage interface from MRIs of the knee. *Physics in Medicine and Biology*, 52:1617–1631, 2007.

[Fukunaga *et al.*, 1997] T. Fukunaga, Y. Kawakami, S. Kuno, K. Funato, and S. Fukashiro. Muscle architecture and function in humans. *Journal of biomechanics*, 30(5):457–463, 1997.

[Galban *et al.*, 2005] C. J. Galban, S. Maderwald, K. Uffmann, and M. E.Ladd. A diffusion tensor imaging analysis of gender differences in water diffusivity within human skeletal muscle. *NMR in Biomed*, 18:489–498, 2005.

[Gea *et al.*, 2006] J. Gea, E. Barreiro, and M. Orozco-Levi. *Skeletal Muscle Plasticity in Health and Disease*, volume 2 of *Advances in Muscle Research*, chapter Skeletal Muscle Adaptations to Disease States, pages 315–360. Springer Netherlands, 2006.

[Gilles *et al.*, 2006] B. Gilles, L. Moccozet, and N. Magnenat-Thalmann. Anatomical modelling of the musculoskeletal system from mri. *Proceedings of MICCAI'06: International Conference on Medical Image Computing and Computer Assisted Intervention*, 4190:289–296, 2006.

[Glocker *et al.*, 2008] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios. Dense image registration through mrfs and efficient linear programming. *Medical Image Analysis*, 12(6):731–741, 2008.

[Golland *et al.*, 2000] P. Golland, W. E. L. Grimson, M. E. Shenton, and R. Kikinis. Small sample size learning for shape analysis of anatomical structures. *Lecture Notes in Computer Science*, pages 72–82, 2000.

[Goodall, 1991] C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 285–339, 1991.

[Gourret *et al.*, 1989] J. P. Gourret, N. Magnenat-Thalmann, and D. Thalmann. Simulation of object and human skin formations in a grasping task. In *SIGGRAPH*, pages 21–30, 1989.

[Gower, 1975] J. C. Gower. Generalized procrustes analysis. *Psychometrika*, 40(1):33–51, 1975.

[Grady and Funka-Lea, 2004] L. Grady and G. Funka-Lea. Multi-label image segmentation for medical applications based on graph-theoretic electrical potentials. *Lecture Notes in Computer Science*, pages 230–245, 2004.

[Gray, 2008] H. Gray. *Gray's anatomy : the anatomical basis of clinical practice*. Edinburgh : Churchill Livingstone„ 40th edition, 2008.

[Harman, 1976] H.H. Harman. *Modern factor analysis*. University of Chicago Press, 1976.

[Heaton and Silverman, 2008] T. J. Heaton and B.W. Silverman. A wavelet-or lifting-scheme-based imputation method. *Journal Royal Statistical Society. Series B Statistical Methodology*, 70(3):567, 2008.

[Heimann and Meinzer, 2009] T. Heimann and H. P. Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, 2009.

[Herzog, 2000] W. Herzog. *Skeletal muscle mechanics: from mechanisms to function*. Wiley & Sons, 2000.

[Hill *et al.*, 1992] A. Hill, T. F. Cootes, and C. J. Taylor. A generic system for image interpretation using flexible templates. In *Proc. BMVC*, pages 276–285. Citeseer, 1992.

[Hill *et al.*, 2000] A. Hill, C. J. Taylor, and A. D. Brett. A framework for automatic landmark identification using a new method of nonrigid correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3):241–251, 2000.

[Hotelling, 1936] H. Hotelling. Relations between two sets of variates. *Biometrika*, 28(3):321–377, 1936.

[Hou *et al.*, 2001] X. W. Hou, S. Z. Li, H. J. Zhang, and Q. S. Cheng. Direct appearance models. In *Proceedings CVPR'01: IEEE Computer Society Conference on Computer Visionand Pattern Recognition*, volume 1, 2001.

[Hyvärinen and Oja, 2000] A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430, 2000.

[Johnson *et al.*, 1998] N. Johnson, A. Galata, and D. Hogg. The acquisition and use of interaction behaviour models. In *Proceedings of CVPR'98: IEEE Computer Vision and Pattern Recognition*, pages 866–871, 1998.

[Jolly *et al.*, 2001] M.P. Jolly, N. Duta, and G. Funka-Lea. Segmentation of the left ventricle in cardiac MR images. In *Proceedings of ICCV'01 :IEEE International Conference in Computer Vision*, volume 1, 2001.

[Jolly, 2006] M.P. Jolly. Automatic segmentation of the left ventricle in cardiac mr and ct images. *International Journal of Computer Vision*, 70(2):151–163, 2006.

[Kaiser, 1958] H. Kaiser. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23:187–200, 1958.

[Kass *et al.*, 1988] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.

[Kaus *et al.*, 2004] M. R. Kaus, J. Berg, J. Weese, W. Niessen, and V. Pekar. Automated segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*, 8(3):245–254, 2004.

[Kelemen *et al.*, 1999] A. Kelemen, G. Székely, and G. Gerig. Three-dimensional model-based segmentation of brain mri. *IEEE Transactions on Medical Imaging*, 18(10):828–839, 1999.

[Kondor and Lafferty, 2002] R. I. Kondor and J. Lafferty. Diffusion kernels on graphs and other discrete structures. In *International Conference on Machine Learning*, pages 315–322, 2002.

[Koryak, 2008] Y. A. Koryak. Functional and clinical significance of the architecture of human skeletal muscles. *Human Physiology*, 34(4):482–492, 2008.

[Kotcheff and Taylor, 1998] A. C. W. Kotcheff and C. J. Taylor. Automatic construction of eigenshape models by direct optimization. *Medical Image Analysis*, 2(4):303–314, 1998.

[Lafon and Lee, 2006] S. Lafon and A. B. Lee. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1393, 2006.

[Lamecker *et al.*, 2004] H. Lamecker, M. Seebass, H.C. Hege, and P. Deuffhard. A 3d statistical shape model of the pelvic bone for segmentation. In *Proceedings SPIE Medical Imaging*, pages 1341–1351, 2004.

[Langs and Paragios, 2008] G. Langs and N. Paragios. Modeling the structure of multivariate manifolds: Shape maps. In *Proceedings CVPR'08: IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[Langs *et al.*, 2006] G. Langs, P. Peloschek, R. Donner, M. Reiter, and H. Bischof. Active feature models. In *Proceedings of ICPR'06: International Conference on Pattern Recognition*, volume 1, pages 417–420, 2006.

[Langs *et al.*, 2007] G. Langs, R. Donner, P. Peloschek, and H. Bischof. Robust autonomous model learning from 2d and 3d data sets. In *Proceedings of MICCAI'07: International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 968–976, 2007.

[Larsen and Hilger, 2003] R. Larsen and K. B Hilger. Statistical shape analysis using non-Euclidean metrics. *Medical Image Analysis*, 7(4):417–423, 2003.

[Leung and Bosch, 2007a] K. Y. E. Leung and J. G. Bosch. Local Wall-Motion Classification in Echocardiograms Using Shape Models and Orthomax Rotations. *Lecture Notes in Computer Science*, 4466:1, 2007.

[Leung and Bosch, 2007b] K. Y. Esther Leung and J. G. Bosch. Localized shape variations for classifying wall motion in echocardiograms. In *Proceedings of MICCAI'07: International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 52–59, 2007.

[Luo and Hancock, 2002] B. Luo and E. R. Hancock. Iterative procrustes alignment with the em algorithm. *Image and Vision Computing - Elsevier*, 2002.

[Macovski, 1983] A. Macovski. *Medical imaging systems*. Prentice-Hall, New Jersey,, 1983.

[Mallat, 1989] S. Mallat. A theory of multiresolution signal decomposition: The wavelet representation. *TPAMI: IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693, 1989.

[Matthews and Baker, 2004] I. Matthews and S. Baker. Active appearance models revisited. *International Journal of Computer Vision*, 60(2):135–164, November 2004. In Press.

[McInerney and Terzopoulos, 1995] T. McInerney and D. Terzopoulos. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4d image analysis. In *Computerized Medical Imaging and Graphics*, volume 19(1), pages 69–83, 1995.

[McInerney and Terzopoulos, 1996] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. In *Medical Image Analysis*, pages 91–108, 1996.

[Melzer *et al.*, 2003] T. Melzer, M. Reiter, and H. Bischof. Appearance models based on kernel canonical correlation analysis. *Pattern recognition*, 36(9):1961–1971, 2003.

[Meyer, 1993] Y. Meyer. Wavelets - algorithms and applications. *Applied Mathematics*, 1993.

[Mikolajczyk and Schmid, 2005] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 27(10):1615–1630, 2005.

[Mitchell *et al.*, 2001] S.C. Mitchell, B.P.F. Lelieveldt, R.J. van der Geest, H.G. Bosch, JHC Reiver, and M. Sonka. Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images. *IEEE Transactions on Medical Imaging*, 20(5):415–423, 2001.

[Mitchell *et al.*, 2002] S. C Mitchell, J. G Bosch, B. P. F Lelieveldt, R. J Van der Geest, J. H. C Reiber, and M. Sonka. 3-d active appearance models: segmentation of cardiac mr and ultrasound images. *IEEE Transactions on Medical Imaging*, 21(9):1167–1178, 2002.

[Movellan, 2002] J. R. Movellan. Tutorial on gabor filters. *Open Source Document*, 2002.

[Nain *et al.*, 2007] D. Nain, S. Haker, A. Bobick, and A. Tannenbaum. Multiscale 3-d shape representation and segmentation using spherical wavelets. *IEEE Transactions on Medical Imaging*, 26(4):598–618, 2007.

[Nain, 2006] D. Nain. *Scale-Based Decomposable Shape Representation for Medical Image Segmentation and Shape Analyses*. PhD thesis, Geoagia Institute of Technology, 2006.

[Narici and Maganaris, 2006] M. Narici and C. Maganaris. *Skeletal Muscle Plasticity in Health and Disease*, volume 2, chapter Muscle Architecture and Adaptations to Functional Requirements, pages 256–288. Springer Netherlands, 2006.

[Neji *et al.*, 2009] R. Neji, A. Besbes, N. Komodakis, J. F. Deux, M. Maatouk, A. Rahmouni, G. Bassez, G. Fleury, and N. Paragios. Clustering of the human skeletal muscle fibers using linear programming and angular hilbertian metrics. In *Proceedings of IPMI'09: Information Processing in Medical Imaging*, pages 14–25, 2009.

[Nowak, 1999] R. D. Nowak. Wavelet-based rician noise removal for magnetic resonance imaging. *IEEE Transactions on Image Processing*, 8:1408–1419, 1999.

[O'Donnell *et al.*, 1998] T. O'Donnell, M.P. Dubuisson-Jolly, and A. Gupta. A Cooperative Framework for Segmentation using Contours and 3D Hybrid Models as Applied to Branching Cylindrical Structures. In *Proceedings of ICCV'98: International Conference on Computer Vision*, page 454, 1998.

[Paragios, 2002] N. Paragios. A variational approach for the segmentation of the left ventricle in cardiac image analysis. *International Journal of Computer Vision*, 50(3):345–362, 2002.

[Pizer *et al.*, 2003] S. M. Pizer, P. T. Fletcher, S. Joshi, A. Thall, J. Z. Chen, Y. Fridman, D. S. Fritsch, A. G. Gash, J. M. Glotzer, M. R. Jiroutek, et al. Deformable m-reps for 3d medical image segmentation. *International Journal of Computer Vision*, 55(2):85–106, 2003.

[Qian *et al.*, 2005] Z. Qian, D. N. Metaxas, and L. Axel. A learning framework for the automatic and accurate segmentation of cardiac tagged mri images. In *Computer Vision for Biomedical Image Applications*, pages 93–102, 2005.

[Rieger *et al.*, 2004] B. Rieger, F. J. Timmermans, L. J. van Vliet, and P. W. Verbeek. On curvature estimation of iso surfaces in 3d gray-value images and the computation of shape descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8):1088–1094, 2004.

[Rissanen, 1978] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465–471, 1978.

[Rogers and Graham, 2001] M. Rogers and J. Graham. Structured point distribution models:. In *Proceedings of BMVC'01: British Machine Vison Conference*, pages 33–42, 2001.

[Rogers and Graham, 2002] M. Rogers and J. Graham. Robust active shape model search. *Lecture Notes in Computer Science*, pages 517–530, 2002.

[Rousson and Cremers, 2005] M. Rousson and D. Cremers. Efficient kernel density estimation of shape and intensity priors for level set segmentation. In *Proceedings of MICCAI'05: International Conference on Medical Image Computing and Computer Assisted Intervention*, volume 1, pages 757–764, 2005.

[Roweis and Saul, 2000] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000.

[Rueckert *et al.*, 2003] D. Rueckert, A.F. Frangi, and J.A. Schnabel. Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. *IEEE Transactions on Medical Imaging*, 22(8):1014–1025, 2003.

[Schclar, 2008] A. Schclar. A diffusion framework for dimensionality reduction. In *Soft Computing for Knowledge Discovery and Data Mining*, pages 315–325. Springer, 2008.

[Schneider, 2001] T. Schneider. Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *Journal of Climate*, 14:853 – 871, 2001.

[Schröder and Sweldens, 1995] P. Schröder and W. Sweldens. Spherical wavelets: Efficiently representing functions on the sphere. In *Proceedings of SIGGRAPH'95*, pages 161–172. ACM New York, NY, USA, 1995.

[Scott *et al.*, 2003] I. M. Scott, T.F. Cootes, and C.J. Taylor. Improving active appearance mode matching using local image structure. In *Proceedings of IPMI'03: Information Processing in Medical Imaging*, pages 258–269, 2003.

[Shi and Malik, 2000] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[Sjostrand *et al.*, 2007] K. Sjostrand, E. Rostrup, C. Ryberg, R. Larsen, C. Studholme, H. Baezner, J. Ferro, F. Fazekas, L. Pantoni, D. Inzitari, et al. Sparse decomposition and modeling of anatomical shape variation. *IEEE Transactions on Medical Imaging*, 26(12):1625–1635, 2007.

[Sjöstrand, 2007] K. Sjöstrand. *Regularized Statistical Analysis of Anatomy*. PhD thesis, Technical University of Denmark, 2007.

[Staib and Duncan, 1992] L. H. Staib and J. S. Duncan. Boundary finding with parametrically deformable models. *TPAMI: IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1061–1075, 1992.

[Staib and Duncan, 1996] L. H. Staib and J. S. Duncan. Model based deformable surface finding for medical images. *IEEE Transactions on Medical Imaging*, 15:720–731, 1996.

[Stegmann *et al.*, 2004] M.B. Stegmann, S. Forchhammer, and T.F. Cootes. Wavelet enhanced appearance modelling. In *International Symposium on Medical Imaging*, pages 1823–1832. Citeseer, 2004.

[Stegmann *et al.*, 2006] M. B. Stegmann, K. Sjostrand, and R. Larsen. Sparse modeling of landmark and texture variability using the orthomax criterion. In *In Proceedings of SPIE'06: International Symposium on Medical Imaging*, volume 6144, 2006.

[Styner *et al.*, 2003] M.A. Styner, K.T. Rajamani, L.P. Nolte, G. Zsemlye, G. Székely, C.J. Taylor, and R.H. Davies. Evaluation of 3D correspondence methods for model building. *Lecture Notes in Computer Science*, pages 63–75, 2003.

[Székely *et al.*, 1996] G. Székely, A. Kelemen, C. Brechbehler, and G. Gerig. Segmentation of 2d and 3d objects from mri volume data using constrained elastic deformations of flexible fourier contour and surface models. *Medical Image Analysis*, 1:19–34, 1996.

[Szeliski *et al.*, 2006] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields. *Lecture Notes in Computer Science*, 3952:16, 2006.

[Taron *et al.*, 2007] M. Taron, N. Paragios, and M-P. Jolly. From uncertainties to statistical model building and segmentation of the left ventricle. In *Proceedings of ICCV'07: International Conference on Computer Vision*, 2007.

[Taron *et al.*, 2009] M. Taron, N. Paragios, and M. P. Jolly. Registration with uncertainties and statistical modeling of shapes with variable metric kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):99–113, 2009.

[Teran *et al.*, 2003] J. Teran, S. Blemker, V. Ng Thow Hing, and R. Fedkiw. Finite volume methods for the simulation of skeletal muscle. *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, pages 68–74, 2003.

[Terzopoulos *et al.*, 1988] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3D shape and nonrigid motion* 1. *Artificial Intelligence*, 36(1):91–123, 1988.

[Tu and Zhu, 2002] Z. Tu and S.C. Zhu. Image segmentation by data-driven markov chain monte carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:657–673, 2002.

[Turkheimer *et al.*, 2006] F. E. Turkheimer, J. A. D. Aston, M. C. Asselin, and R. Hinz. Multi-resolution bayesian regression in pet dynamic studies using wavelets. *Neuroimage*, 32:111–121, 2006.

[Üzümcü *et al.*, 2003] M. Üzümcü, A. F. Frangi, M. Sonka, J. H. C. Reiber, and B. P. F. Lelieveldt. ICA vs. PCA active appearance models: Application to cardiac MR segmentation. In *Proceedings of MICCAI'03: International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 451–458. Springer, 2003.

[van Ginneken *et al.*, 2002] B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever. Active shape model segmentation with optimal features. *IEEE Transactions on Medical Imaging*, 21(8):924–933, 2002.

[Vandergheynst and Wiaux, 2010] P. Vandergheynst and Y. Wiaux. *Four Short Courses in Harmonic Analysis: Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis*, chapter Wavelets on the sphere. 2010.

[Vapnik, 2000] V. N. Vapnik. *The nature of statistical learning theory*. Springer Verlag, 2000.

[Veksler, 2000] O. Veksler. Image segmentation by nested cuts. In *Proceedings CVPR'00: IEEE Computer Society Conference on Computer Visionand Pattern Recognition*, volume 1, 2000.

[Viola and Jones, 2001] P. Viola and M. Jones. Robust real-time object detection. In *International Journal of Computer Vision*, 2001.

[Wallace and Dowe, 1999] C. S. Wallace and D. L. Dowe. Minimum message length and Kolmogorov complexity. *The Computer Journal*, 42(4):270–283, 1999.

[Welch and Witkin, 1994] W. Welch and A. Witkin. Free-form shape design using triangulated surfaces. *Computer Graphics*, page 94, 1994.

[Wolstenholme and Taylor, 1999] C. B. H Wolstenholme and C. J. Taylor. Wavelet compression of active appearance models. *Lecture Notes in Computer Science*, pages 544–554, 1999.

[Xu and Prince, 1998] C. Xu and J.L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing*, 7(3):359–369, 1998.

[Yezzi *et al.*, 1997] A. Jr. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum. A geometric snake model for segmentation of medical imagery. *IEEE Transactions on Medical Imaging*, 16:199–209, 1997.

[Yu *et al.*, 2007] P. Yu, P. E. Grant, Y. Qi, X. Han, F. Segonne, R. Pienaar, E. Busa, J. Pacheco, N. Makris, R. L. Buckner, P. Golland, and B. Fischl. Cortical surface shape analysis based on spherical wavelets. *Transactions on Medical Imaging*, 26(4):582–597, 2007.

[Zhan and Shen, 2006] Y. Zhan and D. Shen. Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method. *IEEE Transactions on Medical Imaging*, 25(3):256–272, 2006.

[Zheng *et al.*, 2007] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features. In *Proceedings of ICCV'07 :IEEE International Conference in Computer Vision*, volume 2007, pages 1–8, 2007.

[Zhu *et al.*, 2009] K. P. Zhu, Y. S. Wong, W. F. Lu, and J. Y. H. Fuh. A diffusion wavelet approach for 3-d model matching. *Computer-Aided Design*, 41(1):28–36, 2009.

[Zollei *et al.*, 2005] L. Zollei, E. Learned-Miller, E. Grimson, and W. Wells. Efficient population registration of 3D data. *Lecture Notes in Computer Science*, 3765:291–301, 2005.

[Zou *et al.*, 2006] H. Zou, T. Hastie, and R. Tibshirani. Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, 15(2):265–286, 2006.